NASA SP-251

SIGNIFICANT ACCOMPLISHMENTS IN SCIENCE AND TECHNOLOGY

Goddard Space Flight Center, 1969

A symposium held at

GODDARD SPACE FLIGHT CENTER

December 3–4, 1969

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

NASA SP–251

SIGNIFICANT ACCOMPLISHMENTS IN SCIENCE AND TECHNOLOGY

GODDARD SPACE FLIGHT CENTER, 1969

The proceedings of a symposium held at the NASA Goddard Space Flight Center, December 3–4, 1969

Prepared by Goddard Space Flight Center

Scientific and Technical Information Division
OFFICE OF TECHNOLOGY UTILIZATION 1970
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
Washington, D.C.

For sale by the National Technical Information Service, Springfield, Virginia 22151 — Price $3.00

FOREWORD

This is an almost verbatim transcript of a Symposium held at Goddard Space Flight Center, Greenbelt, Maryland, on December 3-4, 1969. No attempt has been made to introduce editorial or stylistic uniformity; on the other hand, every attempt has been made to retain the informal flavor of the proceedings.

Page intentionally left blank

CONTENTS

Session I—Science Accomplishments, 1969

Session II—Technology Accomplishments, 1969

I
SCIENCE

Page intentionally left blank

OPENING REMARKS BY DR. JOHN F. CLARK
DIRECTOR OF GODDARD SPACE FLIGHT CENTER

I complement all of you on your discrimination in choosing to listen to the results of our space program today rather than to engage in the sometimes frustrating task of managing the NASA space flight projects.

I also thank you for permitting me to enjoy a couple of days of the same kind of menu. We do this only once a year because we do not want to spoil ourselves.

This is the second of the annual GSFC symposia on the achievements in science and technology in the past year. It really is not quite fair to call it a second symposium on science and technology because last year's meeting was restricted essentially to our science results. This year it has been expanded by what might be called popular demand.

Those of you who were here last fall will recognize that we are doing something a bit different this year. Rather than having a fairly formalized review of our Advanced Research and Technology/Supporting Research and Technology (ART/SRT) program by task area and by organization, we are now working this into an overall review of our science and technology program. Specifically, we are incorporating results from the many space flight programs underway at GSFC.

I think the reason for this is clear. With the maturing of the program, many of the results from the space flight programs are based on earlier ART/SRT results. In turn, many of these space flight program results will be the source for many of the intriguing new ART/SRT programs of tomorrow.

Obviously, these things are all interrelated, but, for the other 363 days of the year, we seem to concentrate our management attention very heavily on the space flight program. Today is certainly an appropriate time to find out that these successful ART/SRT programs are the cornerstone on which the successful space flight program is based.

With that, I would like to introduce our Director of Space and Applications Sciences, chairman of today's session, Dr. George Pieper.

REMARKS BY DR. GEORGE F. PIEPER, CHAIRMAN

Today, members of the Goddard Space Flight Center are pleased to present to you a review of their accomplishments in science thus far this year, 1969. The format for today's program will be the same as in our review for 1968. Each speaker is allotted five minutes and three slides.

Experience with the earlier program and with the preparation for this one confirms our belief that the method is good for both the speaker and the audience. It forces the speaker to select very carefully the point he wants to make, and it leaves the audience secure in the knowledge that the subject will be changed in five minutes.

In order to keep the program on schedule, I shall limit questions from the floor to no more than one or two short ones after each talk. Our speakers will be pleased to talk with you at the coffee breaks, lunch, or at the reception in the library at the end of the first day.

MULTISPECTRAL TERRAIN PHOTOGRAPHY FROM APOLLO 9

Paul Lowman

CHAIRMAN:

I would like to begin the program by calling on our first speaker. Dr. Paul Lowman has been the principal investigator in multispectral terrain photography from several Apollo flights. He will tell us about his work on Apollo 9.

DR. LOWMAN:

I think most of you are familiar with the terrain photography experiments we carried on the Mercury, Gemini, and Apollo 7 flights, in which we used single, hand-held cameras to photograph selected parts of the Earth's surface for geology, geography, and oceanography.

On Apollo 9 we carried out a much more sophisticated experiment in which we used four coaxially mounted cameras simultaneously, photographing the Earth in an attempt to prove, or demonstrate, the value and feasibility of multispectral terrain photography for Earth resources purposes.

Figure 1 shows the cameras we used on Apollo 9; they are mounted in the hatch window of the command module. We used four Hasselblad 500 E.L.'s, the electrically driven model similar to the manual model we used successfully on Gemini and Apollo 7.

The critical experiment involved using the cameras with different film-filter combinations. We had two cameras with black and white film, using green and red filters, one camera with color infrared film, and one with black and white infrared film. With each click of the shutter, we got a set of four pictures, as shown in Figures 2 through 5.

The area shown is in northern Alabama, just east of Birmingham, and includes the southern part of the Coosa River Valley and Ridge Province, part of the Blue Ridge Province, and part of the Piedmont Plateau. As is typical of terrain photography taken from this altitude, each picture covers an area extending about 70 miles on a side.

Every one of these films has a band-pass with its own particular uses. Figure 2 is the black and white infrared photograph. The nice feature of this film is that bodies of open water stand out very sharply. The Coosa

River, which has been dammed for reservoirs in several places, is clearly revealed. The Blue Ridge Province stands out in the center of this figure. Figures 3 and 4 are from black and white film using the red and green filters respectively. The black and white with the red filter is good for topographical mapping because of its high resolution. Birmingham is visible as a white region on the extreme left of Figure 3. Figure 4, taken with the green filter, may not look very impressive, but it is pretty good in combination with the other two black and white bands.

Figure 5 was taken with the color infrared film. This is probably the best general-purpose picture, as can be seen. Bodies of water stand out as blue, and the vegetation appears red. One can see the vegetation and land-use patterns very clearly. Because the geology of this area is expressed by vegetation patterns, especially in the valley and Ridge Province, we can see the geological structure more clearly. The Piedmont now appears as the darker region on the right of the figure.

Now I shall discuss another area we have been studying using a picture taken with the color infrared film. The area shown in Figure 6 is southwest California, including San Diego, with the Mexican border near the bottom of the figure. Once again this is a large area, stretching about 70 miles on a side.

The particular problem I have been studying is the regional structure of the Peninsular Ranges, which are the mountains along the coast, seen in the top half of the figure. Visible is a series of major northwest- and southeast-trending faults. The San Andreas Fault is just barely visible in the top right corner, and the San Jacinto Fault can be seen on the top right side. I have been working with the Elsinore Fault, named from Elsinore Lake, seen at the very top of the figure. The Elsinore Fault extends down to and supposedly through the Aqua Tibia Mountains, past Lake Henshaw (seen slightly above center), and down all the way to the Mexican border. These northwesterly trending faults are generally supposed to be strike-slip, or wrench, faults, that is, faults with primarily horizontal movement. The particular problem is: How much horizontal movement has there been on the Elsinore Fault, or, in fact, has there been any at all?

In the figure we first found that we could see a number of northeasterly trending faults which cross the Elsinore Fault without being noticeably displaced. This indicates that there has not been any major lateral movement along the Elsinore Fault, as far as we can see.

Another thing we can see from the figure is that the Elsinore Fault comes down to the north end of the Aqua Tibia Mountains, whereupon it

diverges around them to the north, rather than going straight through. This again argues against major lateral movement along the Elsinore Fault as a whole.

The northeasterly trending faults, near the bottom of Figure 6, are not shown on the existing geologic maps. To some extent, then, they have been discovered by the Apollo 9 photography. Of course, these are probably not active faults. This is a very quiet area seismically, but these faults are nevertheless important. In reservoir construction, for example, flooding any of these faults and fault-controlled valleys requires knowledge of any fault controls to prevent possible reactivation of the faults, which would cause earthquakes, either by the load of water in the reservoirs or by lubrication.

In summary, I think we can say that we have thrown some fundamental new light on the geologic structure of southern California.

MEMBER OF THE AUDIENCE:

Was the equipment that you had for ground data handling adequate for the purposes of your analysis work?

DR. LOWMAN:

No, it was not. I have made only the rather crude analysis of this photography. One problem in particular we have had with this kind of work is that you have to get out into the field; and of course this takes quite a lot of time. In addition, we did not have as much equipment as we should have had to study these photographs. For example, we would like to be able to recombine the three black and white bands. You can see I used chiefly the color infrared film. We should be able to recombine those three black and white bands as well.

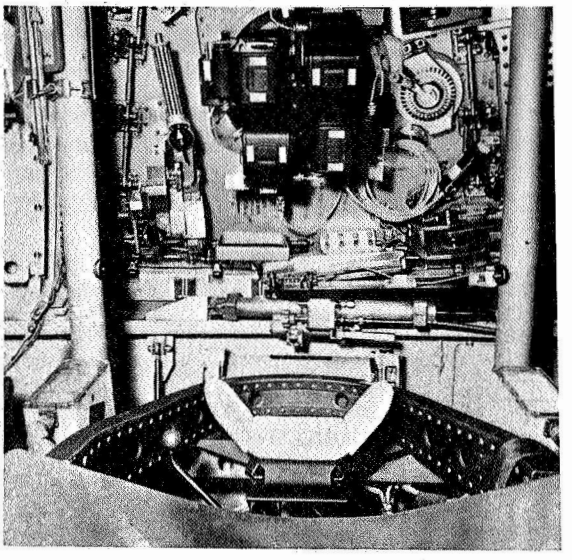


Figure 1

Figure 2

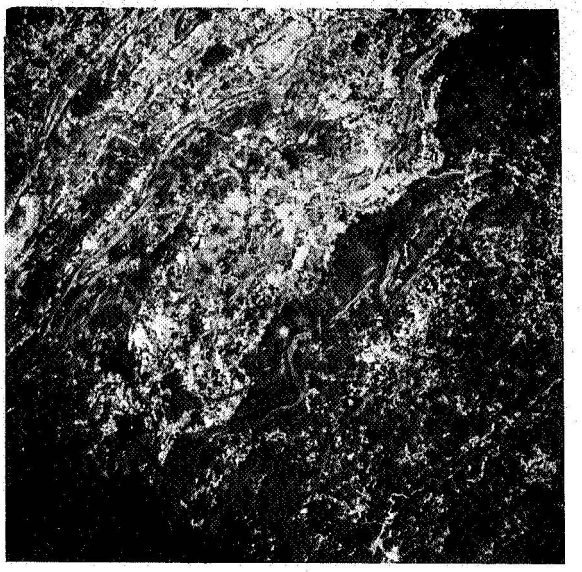


Figure 3

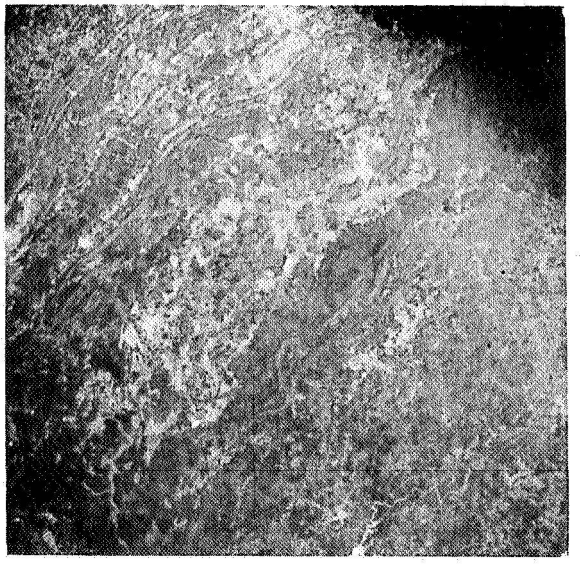


Figure 4

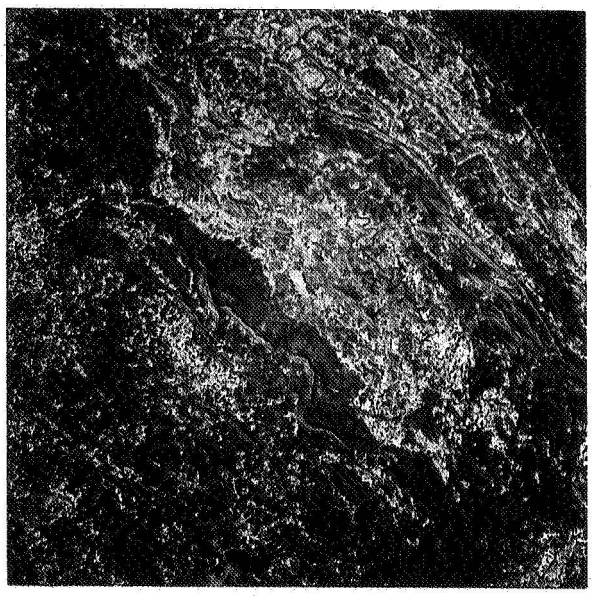


Figure 5

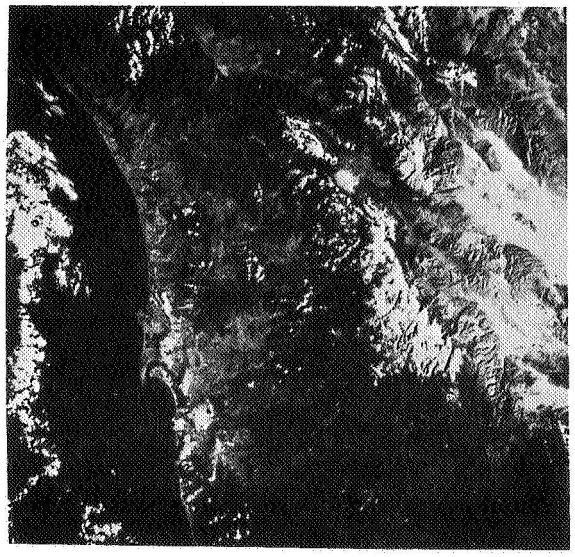


Figure 6

REMOTE SENSING OF GLOBAL OZONE FROM NIMBUS 3

Cuddapah Prabhakara

CHAIRMAN:

Next I will call on Dr. Cuddapah Prabhakara who will talk about global ozone measurements made from the Nimbus 3 satellite.

DR. PRABHAKARA:

Ozone is a minor constituent of the Earth's atmosphere. It is produced by the interaction of solar near-ultraviolet radiation with the air. Principally this gas is present in the stratosphere between the heights 10 to 40 km. The average amount of this gas, in a vertical column of unit cross section, is about 0.35 cm (STP). In spite of such a small amount, ozone is primarily responsible for heating the stratosphere by absorbing the solar ultraviolet energy. Further, ozone acts as a useful tracer gas to follow the air motions in the atmosphere. Ozone variations in association with cyclonic and anti-cyclonic flow patterns in the atmosphere could prove to be helpful in predicting the weather.

Owing to these several facts, a detailed study of atmospheric ozone on a global basis is necessary and potentially valuable. With the advent of the polar orbiting Nimbus 3 satellite, it is now possible for the first time to do such a global study of atmospheric ozone.

On board the Nimbus 3 satellite, a high resolution infrared interferometer spectrometer (IRIS) measured the infrared spectrum emerging from the Earth and its atmosphere. In Figure 1 we show a sample spectrum taken over Brownsville, Texas. The 9.6μ region of such a spectrum contains information about the atmospheric ozone. In addition to this band, the 15μ band can help us determine temperature distribution while the 6.3μ band yields the water vapor distribution.

We have inverted these 9.6μ radiance measurements and obtained the vertical distribution and the total amount of ozone in a vertical column. In the method of inversion, the vertical ozone profile is characterized by a known climatological mean and an unknown deviation from this mean. We solve for this deviation using a nonlinear iterative scheme to get an optimum solution.

In Figure 2 we see one such solution obtained by inverting the IRIS spectrum taken over Point Mugu, California, on the 28th of June 1969. In order to verify our solution, we compare the vertical sounding of ozone made by a balloon-borne ozonesonde which was flown to coincide with the overpass of Nimbus 3 at Point Mugu. The ozone profile derived from the IRIS

spectrum agrees well with that observed by the balloon instrument. This example illustrates the reliability of our method.

We have applied this remote sensing technique to the data taken over the globe and derived a global distribution of total ozone for the 22nd of April 1969 (see Figure 3). We find several striking ozone maxima and minima. These ozone features are intimately related to the atmospheric circulation.

From conventional meteorological maps of the Northern Hemisphere, we find that the ozone maximum over the United States is associated with a low pressure system. A similar situation exists over Siberia. In the tropics over Africa, the ozone minimum is associated with an anticyclone in the upper air. At the present time, conventional meteorological maps of the Southern Hemisphere are lacking, but from the ozone maxima shown in our map we can locate the low pressure systems in the Southern Hemisphere.

The large scale behavior of atmospheric ozone as revealed in this study can be usefully applied to the prediction of global weather.

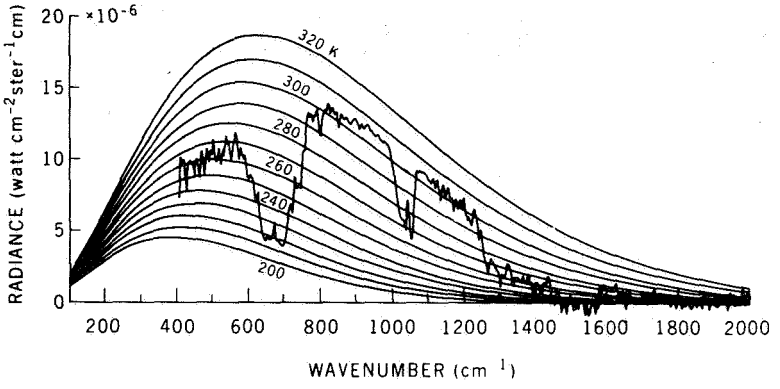


Figure 1 – Nimbus 3 iris spectrum over Brownsville, Texas (27°N), 22 April 1969 (R. A. Hanel).

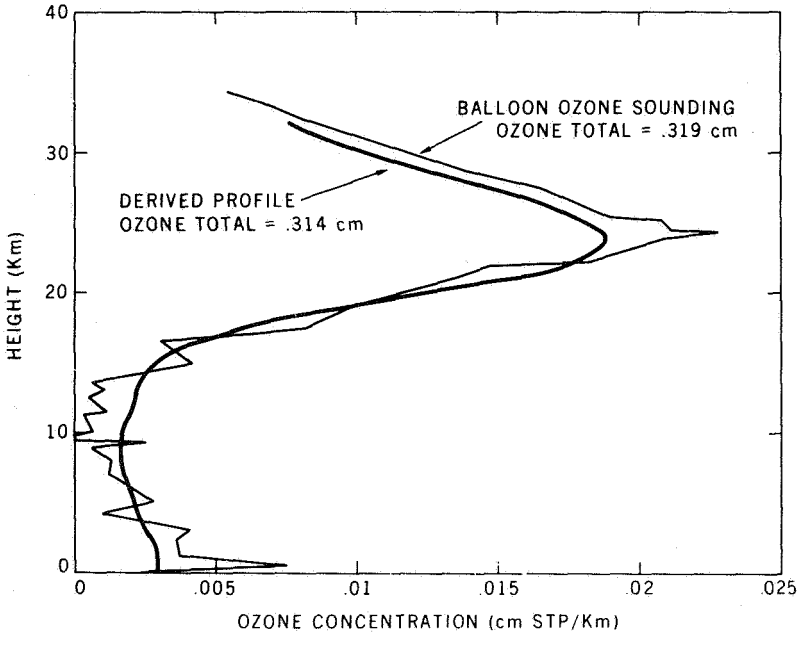


Figure 2 – Nimbus 3, orbit 1012 over Point Mugu, California, 28 June 1969.

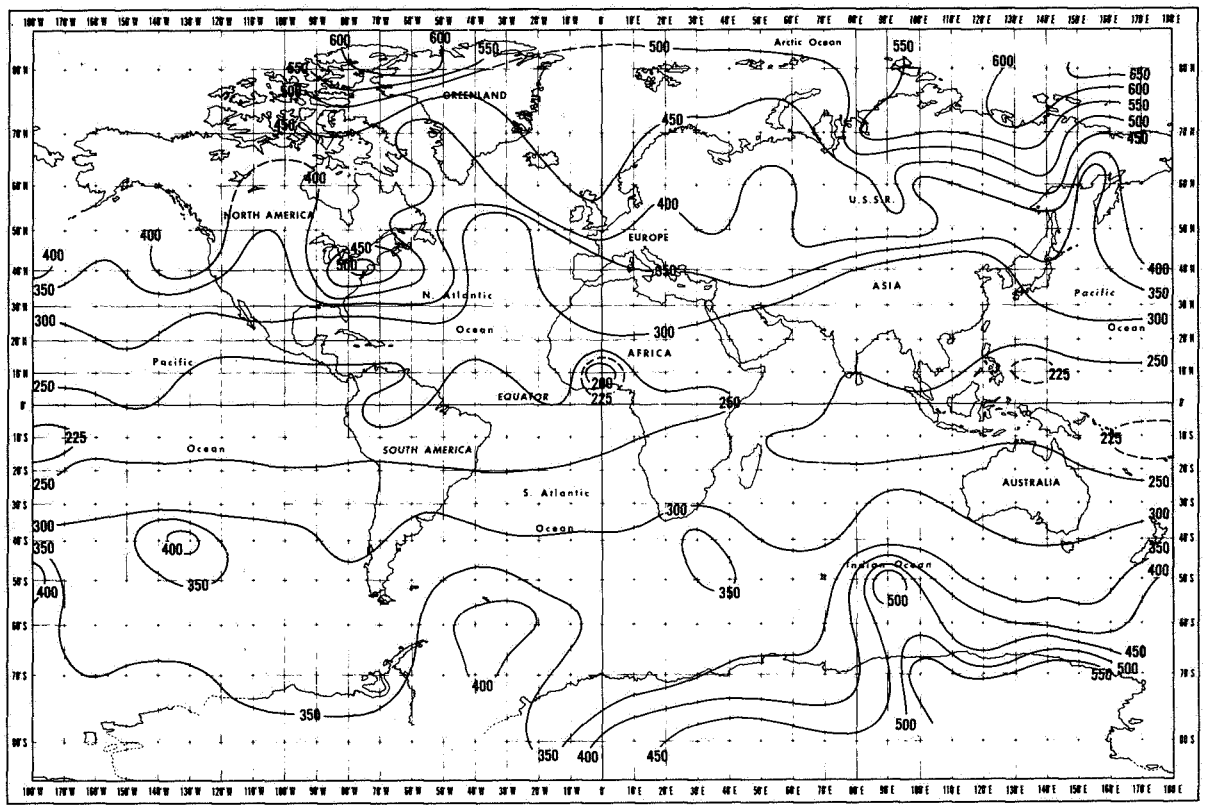


Figure 3 – Total ozone count (10^{-3} cm STP) from Nimbus 3, 22 April 1969.

MRIR OBSERVATIONS OF JET STREAMS

Vincent V. Salomonson

CHAIRMAN:

Our next speaker is Dr. Vincent Salomonson, who will talk about medium resolution infrared observations of jet streams.

DR. SALOMONSON:

Since the launching of the first meteorological satellite, there has been considerable interest in using observations from these satellites to locate and study jet streams. It is of particular interest to see if jet streams and attendant flow conditions can be observed by radiometers responding to far infrared radiation because these radiometric measurements provide meaningful observations during both daytime and nighttime hours when clouds may or may not be present.

A study has recently been accomplished in which Nimbus 2 medium resolution infrared radiometer (MRIR) observations of 60 cases of southwesterly subtropical jet stream flow occurring over or near the continental United States were examined. The specific objectives of this study were to provide quantitative answers to two questions: Do consistent patterns in equivalent blackbody temperature occur relative to jet stream speed maxima, and are the magnitudes of these temperatures correlated with the maximum wind speed and a measure of the large scale lateral shear?

Equivalent blackbody temperature fields associated with each of the 60 jet stream cases were analyzed on Mercator map base with an approximate scale factor of 1:10 million. These temperature fields represented the emitted radiation observed in the 6.4-6.9, 10-11, and 5-30 micron spectral regions. A diamond-shaped grid, shown in Figure 1, consisting of points spaced 5° of longitude apart, was used to obtain a sample of temperatures from each jet stream case at 13 locations relative to the wind speed maximum existing at the center of the grid. The grid was oriented so that its tangential axis was parallel to the flow passing through the wind speed maximum.

It was found that characteristic patterns are observed in the vicinity of speed maxima occurring in subtropical jet stream flow. A meaningful pattern can be most consistently seen in the 6.4-6.9 micron mean equivalent blackbody temperature displays shown in Figure 2. The predominant feature of these patterns is the band of warmer temperatures to the rear of the wind

speed maximum and to the left of the jet stream core. The observed location of this band of warmer temperatures is in agreement with vorticity advection considerations that call for subsiding and drier air in this area, thereby resulting in warmer equivalent blackbody temperatures. The wind speed maximum occurs, in the mean, near the 240° K isotherm.

A stepwise multiple regression procedure was used to ascertain the amount of correlation between the magnitudes of the equivalent blackbody temperatures at specific grid locations and maximum wind speed and wind shear. In dependent and independent sample tests, the best overall results were obtained with a regression relationship involving 5-30 micron predictors from four of the 13 possible grid locations. This relationship is shown in Figure 3, with the locations involved schematically illustrated at the bottom of the figure. The multiple correlation coefficients obtained in dependent and independent tests were 0.71 and 0.65 respectively. Both results were significant at the 95% level. Equivalent blackbody temperatures were found to be much less related to wind shear largely because of greater errors inherent in the determination of the wind shear magnitudes.

It was concluded from this study that patterns of equivalent blackbody temperature obtained from MRIR observations can be used to locate regions of southwesterly jet stream flow in the subtropics and to estimate the position and magnitude of the maximum wind speed. The results suggest that further improvement and extension of the results to a greater variety of jet stream occurrences may permit the quantitative use of these MRIR observations in studying the role of the jet stream in the general atmospheric circulation and in delineating areas of clear air turbulence.

MEMBER OF THE AUDIENCE:

To what altitudes do those equivalent temperatures refer, and how did you get your wind speeds?

DR. SALOMONSON:

The 6.4-6.9 micron results are most related to the temperature and humidity of the atmosphere near the 400-millibar level. The estimates of wind speed were obtained from the National Meteorological Center (NMC) maximum wind speed analyses. The NMC wind shear analyses and the 200- and 300-millibar upper-air analyses were also used in determining the locations of the wind speed maxima.

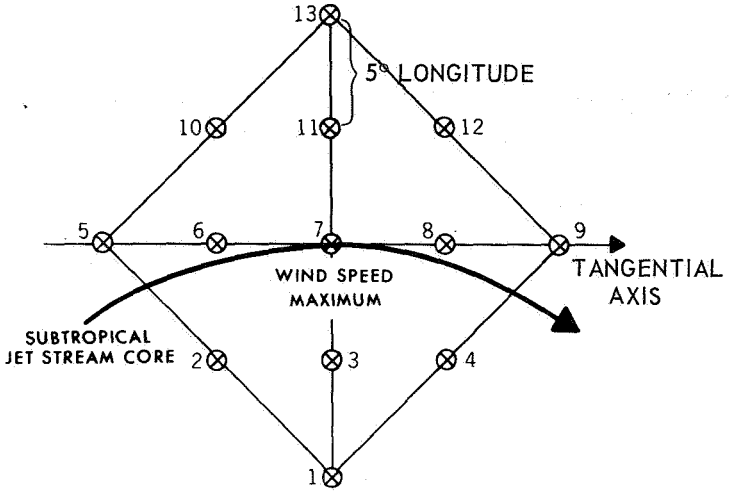


Figure 1 – Grid used for obtaining equivalent blackbody temperatures at 13 points relative to a wind speed maximum.

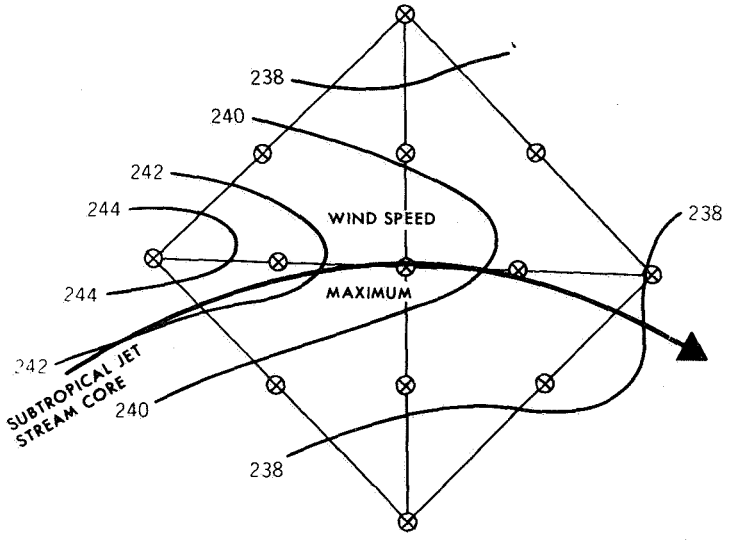


Figure 2 – Mean equivalent blackbody temperature patterns for the 6.4 - 6.9 micron region.

REGRESSION EQUATION FOR TEMPERATURE VERSUS WIND SPEED

$$\tilde{V} = -408.4038 + 0.78643\,T(1) + 0.70356\,T(2)$$

$$+1.07342\,T(12) - 0.65427\,T(13)$$

MULTIPLE CORRELATION COEFFICIENT

R = 0.71 DEPENDENT SAMPLE

R = 0.65 INDEPENDENT SAMPLE

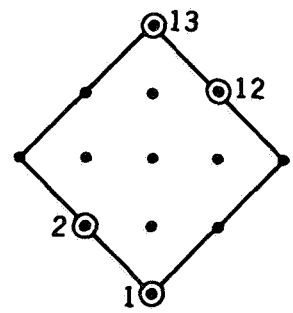


Figure 3

OPTICAL PROPERTIES OF CLOUDS

Warren A. Hovis

CHAIRMAN:

Next, we turn to the subject of clouds. The first talk on clouds will be given by Dr. Warren Hovis.

DR. HOVIS:

The research I would like to describe is devoted to the measurement of the optical properties of clouds and is aimed at developing a method of remote sensing of cloud type–the size of the cloud particles and whether they are water droplets or ice particles–from satellite measurements.

These measurements were all made with a downward-looking spectrometer mounted on various aircraft. We found that the reflected solar infrared region was best for our purposes.

Figure 1 shows the reflectance of five clouds measured between 1.6μ and 2.3μ. The stratus and stratocumulus are water droplet clouds; the cirrus and cirrostratus are ice crystal clouds.

In the region from 2μ to 2.1μ, ice has an absorption band, so one would expect the reflectance to be low, as indeed it is, for the ice crystal cloud. Water clouds vary in reflectance, the variation depending on the water droplet size: the smaller the droplet size the more efficient the scattering and the higher the reflectance.

Centered about 1.9μ, there is an absorption band that is due to water vapor. To the left of this band, we find that all water droplet clouds reflect almost uniformly at about 1.72μ. The ice crystal clouds vary in reflectance with particle size, again the smaller particles reflecting more intensely than the larger particles.

These spectra summarize results from a number of measurements. We have many more. In general, they are true of all naturally formed clouds. They suggest that a two-channel radiometer with bands centered at 1.7μ and 2.1μ mounted on a satellite could be used to discriminate cloud type and determine to some degree the size of cloud particles.

Now I stress that these values are correct for naturally formed clouds because we encountered one anomaly that does not follow the rules, an unnatural cloud, namely a jet contrail.

Jet contrails are becoming more important to our weather. On one particular day (February 18, 1969) when we made flights, there was about a 30% light cirrus cover over the city of Cleveland, and virtually all of that cirrus cover was due to long, persistent jet contrails. We were fortunate in being able to measure the reflectance of one of these contrails, and we compared it with the reflectance of natural clouds.

Figure 2 shows the reflectance of low stratus, cirrus, and the jet contrail. One can see that around 2.1μ to 2.2μ, the jet contrail reflectance is fairly low. The altitude of the contrail was 29,000 ft; the outside air temperature was about $-50°$C. Certainly this was an ice cloud, since it existed for about 15 minutes at these temperatures. However, when viewed in the shorter wavelengths, it was found to be one of the brightest clouds we have ever seen, reflecting far more than the stratus or the cirrus.

This suggests that, as jet contrails become more important to our weather, we might add a third channel centered at 1.6μ to discriminate between jet contrails and natural clouds.

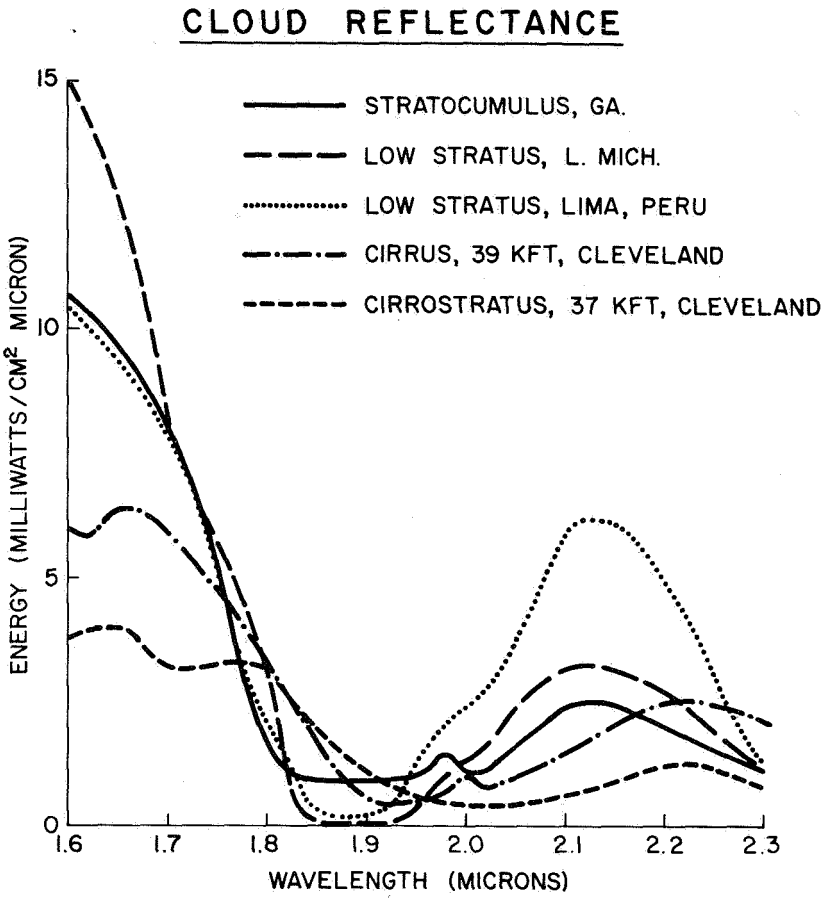

CLOUD REFLECTANCE
STRATOCUMULUS, GA.
LOW STRATUS, L. MICH.
LOW STRATUS, LIMA, PERU
CIRRUS, 39 KFT, CLEVELAND
CIRROSTRATUS, 37 KFT, CLEVELAND
ENERGY (MILLIWATTS/CM² MICRON)
15
10
5
0
1.6
1.7
1.8
1.9
2.0
2.1
2.2
2.3
WAVELENGTH (MICRONS)

Figure 1

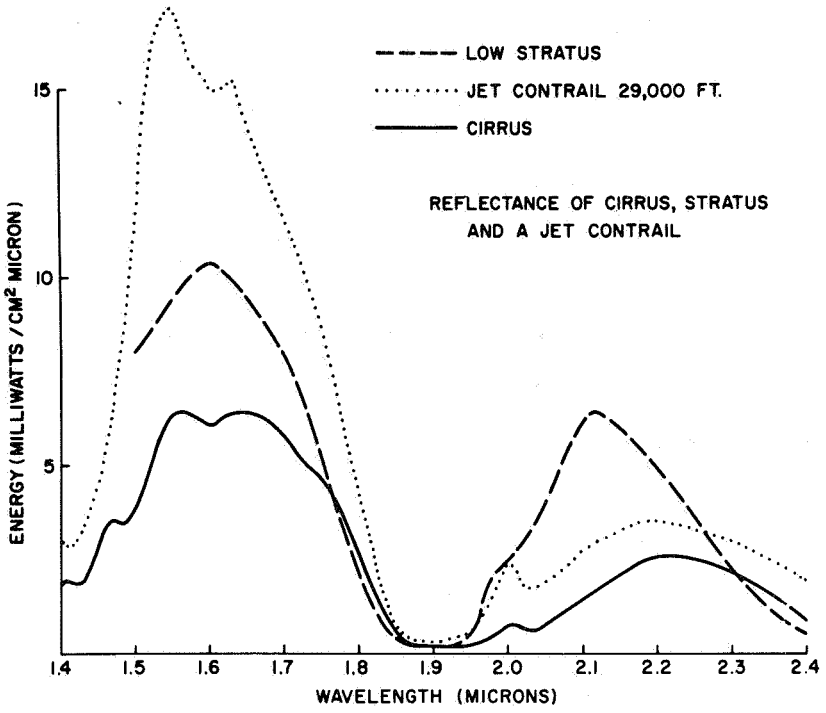

LOW STRATUS
JET CONTRAIL 29,000 FT.
CIRRUS
REFLECTANCE OF CIRRUS, STRATUS
AND A JET CONTRAIL
ENERGY (MILLIWATTS / CM² MICRON)
15
10
5
1.4
1.5
1.6
1.7
1.8
1.9
2.0
2.1
2.2
2.3
2.4
WAVELENGTH (MICRONS)

Figure 2

A COMPARISON BETWEEN OBSERVED WINDS AND CLOUD MOTIONS DERIVED FROM NIMBUS 2 HRIR MEASUREMENTS

William Shenk

CHAIRMAN:

Next then we will move on to Mr. William Shenk, who will talk about cloud motions and wind streams.

MR. SHENK:

Undoubtedly, one of the first measurements that man ever made of the wind speed and wind direction was with an uplifted moistened finger. Fortunately, we have made progress in our wind measurement capability since that time. One of the most successful experiments in this regard was the data acquired from the ATS 1 and ATS 3 geosynchronous satellites, where it was determined that judiciously selected small clouds could serve as representative tracers of the wind. The data from those sensors are particularly useful in the tropics and in the mid-latitudes. In high latitudes, the obliquity of the view perhaps will preclude the use of the measurements for wind determination. Therefore, another technique was investigated where two adjacent orbits from a polar orbiting satellite were used to try to ascertain the cloud motion between the views of the clouds in a 108-minute period.

The particular satellite data selected were those of the high resolution infrared radiometer (HRIR) on board the Nimbus 2 satellite. Thus, not only could we ascertain whether we could determine cloud motion from two views of a cloud, but we could also try to see if a radiometer was useful for this purpose.

Figure 1 gives us an illustration of the particular technique used when we see a cloud at two different times. The first view of the cloud is on the left side of Figure 1, and the second view is on the right side. The portion of the cloud that was used for the motion is shown by the two cross-hatched regions, and the centroid is shown by the circled cross. In 108 minutes the cloud moved the distance shown between the two centroid locations.

An air parcel located at 750 mb had moved the distance indicated by the dashed vector during the 108-minute period. The actual wind speed determined from constant pressure charts was 290° at 35 kt, which compares favorably with the cloud motion of 295° at 40 kt.

The cloud height was determined by using the effective blackbody temperature that was measured by the HRIR. Then, that temperature was located on the constant pressure chart and the level of the cloud top interpolated.

Figure 2 shows the frequency distribution of the cloud top heights that were measured using the entire data sample (43 vectors). The low and middle and the high cloud distributions are rather clearly defined. No effective blackbody temperatures below 240° K were used, due to the methods by which the data were processed. Therefore, the high cloud envelope could very well extend higher than is shown, where one would expect to find more cirrus. However, in the case of the lower and middle clouds, the frequency distributions are probably quite representative of our sample. These frequency distributions are also encouraging from the viewpoint that we think that we have selected clouds that fill the field of view of the radiometer and are reasonably opaque. This is a very important consideration, since the cloud top could be incorrectly located if this were not the case.

Figure 3 shows us the comparison between the computed cloud motions from the HRIR data along the ordinates and the observed wind speeds as determined from the constant pressure charts along the abscissa. One sees that this agreement is fairly good; that is, within generally five or ten knots. The 45° line represents a perfect correlation. A similar curve was determined for wind direction, with similar results.

In conclusion, we have had some success in determining wind speeds from cloud motions using a polar orbiting satellite. In addition, it is also encouraging to see that we can use a radiometer for this purpose.

MEMBER OF THE AUDIENCE:

You note from that last figure that your cloud motion compared systematically is more rapid than your 45° line. Is that significant?

MR. SHENK:

I think when you get up to higher wind speeds it may very well be. This is because you are getting closer to the region of cirrus clouds, where the cirrus are less opaque. Thus, the higher you go, the greater the danger in making a height error. Naturally, if you make the assumption the winds increase with height, then the cloud motions would appear to be faster than the observed winds.

MEMBER OF THE AUDIENCE:

How can you tell the velocity of a cloud relative to the wind?

MR. SHENK:

The cloud motions were determined from the movement of the cloud from one orbit to the next. The wind speeds were determined by using constant pressure charts, after the height had been determined, to then ascertain the speed and direction of the wind at that time at that place.

MEMBER OF THE AUDIENCE:

How do you insure that you have located the same cloud in each orbit?

MR. SHENK:

One of the most effective methods of doing this was to assume that the cloud top temperature was fairly constant with time. And it generally turns out that at high latitudes, where the amount of convective clouds is somewhat less, this is not a bad assumption. If one were working with this technique in the tropics, where a great deal of the clouds were convective, the cloud top temperature would probably change radically in a 108-minute period. The amount of overlap between two orbits of a polar orbiting satellite is not very great in the tropics, with the satellite at a height of 1100 km above the earth.

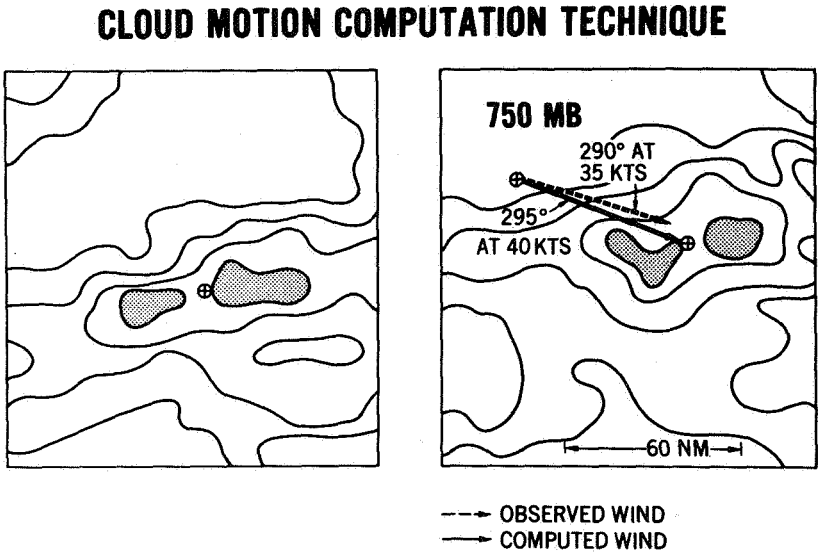


Figure 1

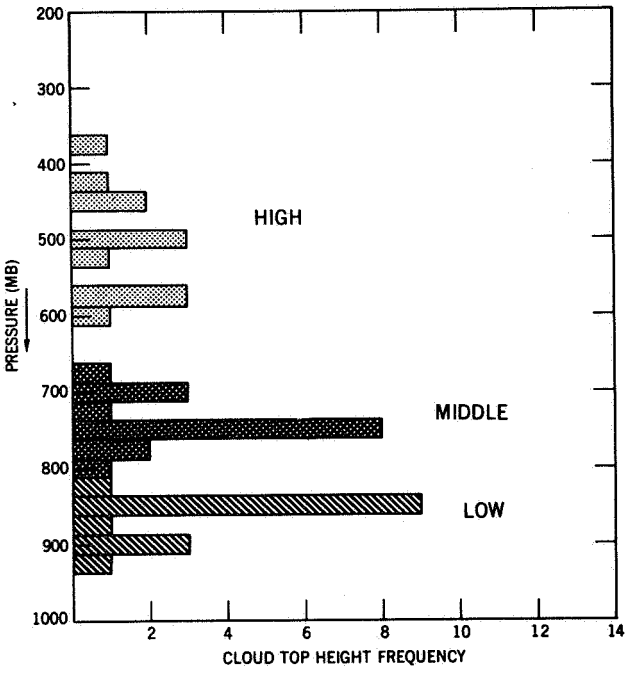


Figure 2

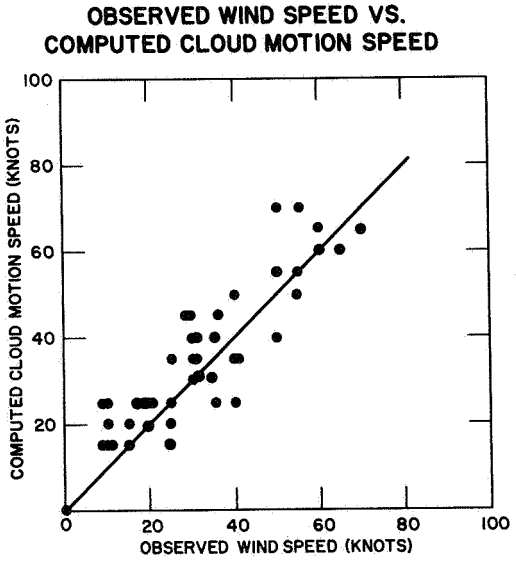


Figure 3

SIMULATED GARP STUDIES

Robert Jastrow

CHAIRMAN:

The Global Atmospheric Research Program (GARP) will be a major undertaking throughout the world and, in particular, in the United States. At the Institute for Space Studies, Dr. Jastrow and his colleagues have been carrying out some simulated GARP studies, which Dr. Jastrow will describe to us now.

DR. JASTROW:

The numerical experiments were started initially by Professor Charney working with us. I deflected them in a direction that concerns the specific applicability of the Nimbus data now being acquired to the simulation studies that are called for in the plan for United States participation in GARP.

We have now gained more fundamental insight into the mechanisms that underlie the results of the experiments that I will now discuss, but I am going to concentrate only on some of their pragmatic aspects. The fundamental work is still being carried on with us by Prof. Charney. These experiments are designed to determine whether winds can be calculated from the actual temperature data obtained with the IR vertical sounders presently on one polar-orbiting satellite, Nimbus 3, and later on two polar-orbiting satellites whose orbits would be spaced 90° apart and which would contain a side-scan capability.

In principle, it is clear that one can determine the winds from the temperature, provided one has an approximation of the initial conditions for the equations of motions and the associated equations that determine the state of the atmosphere. For instance, if one knows the temperature, in both space and time, and the surface pressure, then one has enough information (given the initial approximation) to carry on the calculations indefinitely. That is a point that emerges just by looking at the equations and their dependence on the variables.

Our experiment started from this viewpoint, but something else emerged that was more significant: One need not know the initial conditions completely, not even the surface pressure, if he has the complete history of the temperature in both space and time. This then determines the winds and the surface pressures and determines them with an accuracy well within the limits set by the GARP data requirements.

Thus, knowing the history of one variable determines all the others. This is a fact which, I think, can be proven analytically for sets of linear equations but is not provable for nonlinear equations. We determined empirically that this is the case for the system of atmospheric equations.

If our results are confirmed by later work with better physical models of the atmosphere than the one we used, there will then be considerable significance in it for the GARP plans, since 90% of the data management problem, and probably of the expense for the GARP global experiment, is associated with the cloud tracking (i.e., wind determination by cloud tracking of geostationary cloud images). There are also delicate problems of meteorological judgment in deciding whether a particular cloud is moving with the wind or not. It is something at present that Fujita can do, but probably not a large number of other people, unless trained by him.

In our experiments we used a program developed by Mintz and Arakawa that was originally formulated under Professor Charney's guidance, I believe. It is a program that has good atmospheric physics in it so that it can be used as the basis for the predictability studies for the GARP program. It is a two-level model, available with grid resolutions of 800 km and 400 km.

The first step consists of generating a history tape by carrying on a calculation, starting from any reasonable set of initial conditions, for about 100 days. The history tape was taken by us to represent the correct atmospheric motion (i.e., the motion that the atmosphere would have, to the extent that the Mintz-Arakawa model can represent the Earth's atmosphere).

As a second step, we went back to Day 85 in the model, far enough along so that all the transients associated with our particular choice of initial conditions had disappeared. We then carried the integration forward a second time, but this time, we started from a perturbed state at Day 85. The results of these calculations are shown in Figure 1.

Curve *A* is neither of the individual calculations mentioned but is instead the rms error, or deviation, of the globally averaged winds, (i.e., between the original unperturbed calculations, the history tape, and the perturbed calculations which start at Day 85). The perturbation was obtained by arbitrarily introducing into the temperature field a random error, and this was run for a day or so until the perturbed winds reached an average of about 1 m/sec different from the unperturbed winds.

The perturbations grow rapidly until, at the end of about 20 days, the perturbations (i.e., the mean differences between the perturbed and the

correct, or unperturbed, winds) are 10 m/sec, which is comparable to the mean wind speed. Therefore, the two flow patterns are completely uncorrelated. This is the background for the statement that one can only predict ahead 5 or 10 days at most, even with a set of global initial conditions.

We went back to Day 86, and we put in Nimbus-type temperature data. That is, we took temperature values from the history tape and introduced random errors with a ±1°C uncertainty in an attempt to mimic the probable errors (i.e., the uncertainty in the IR vertical-sounding data). We put them back into the calculation, not at every point or at every moment but with a frequency in time and a spacing on the grid that represented the average coverage of the Nimbus polar-orbiting satellite.

We found that this constrains the calculation so that it must unfold within the constraints of these temperature data and their uncertainties and restores the winds to their correct value (curve *B*), preventing the perturbation from growing. It keeps the flow pattern down to about 1 m/sec for a global average of the correct flow pattern. In other words, it suppresses the effect of the perturbation. Curve *B* indicates that if one starts with reasonable winds, he can then continue to determine the winds without any cloud tracking, simply by using the satellite's vertical-sounding data.

We did another experiment in which we went ahead about 10 days to a point where the winds were 4.5 m/sec from their correct values. That is a rather bad representation of the original state. We started to insert temperature data at that point, and we found that, with some patience—waiting about 10 or 15 days—the temperature data again drove this flow field (curve C) down toward the correct flow field. It can be seen that one does not even have to have a reasonable approximation of the initial winds. If one waits about 10 days, the history of the temperature field will force the flow pattern into the correct one regardless of what is used initially.

Most recently, we repeated this analysis using a system that closely imitates the GARP system—that is, two polar-orbiting satellites, 90° apart, with the data fed in not in an average representation of space-time frequency in the global system but in a manner that closely followed the actual sub-satellite track. We fed the data in as it will be fed into the GARP system—on a real-time basis. Figure 2 shows the results for data closely mimicking the real-time input planned for the GARP global experiment. The results are about the same as those obtained in the earlier, rougher experiment.

We added hypothetical coverage by geostationary satellites to the second experiment, assuming that vertical sounders could be put on such

satellites, with coverage going up to 52° north and south latitude. As can be seen in Figure 2, this does not really improve the results. It does bring the values down quicker to the observed flow, but the final observed flow is no better than it is without the additional satellites.

Figure 3 shows the relation between the mean error in winds and the mean error in temperature. A temperature error of 1°C is the current figure quoted in the GARP plan and that produces an error of about 1 m/sec in the winds, well within the GARP data requirements.

Since there are many vertical soundings taken, three or four to the average grid spacing in the GARP plan, we can, by smoothing and averaging these data, reduce the temperature error by a factor of $\sqrt{3}$ to about 0.5°C and bring the wind error down to about 0.5 m/sec, which is good even in the tropics. The average error in surface pressure corresponding to a 0.5°C temperature error is 0.5 mb. At a 1°C temperature error, the average error in surface pressure is 1 mb, again well within the GARP data requirements.

In conclusion, there are two things that can be said. One is that if these results are confirmed by model experiments with better physical models (something that we would like to do and that other people are doing), the effect will be to place the heaviest part of the burden of the global weather observing system on the satellites. It really means that it will be unnecessary to track clouds. Of course this is an assertion that will have to be checked by comparing the cloud-tracking results with these calculations. However, the entire balloon and buoy system for measuring surface pressures and the cloud tracking that is such a heavy element in the data volume of the GARP system can be discarded if this preliminary result is confirmed by more careful work.

Second, one can use the same kind of simulation to attack the problems of the observing-system simulation experiment that is cited in the plan for United States participation as being "the preliminary experiment prior to the global experiment, of the highest urgency, and essential for the sound design of the global experiment." We hope to start the second round of variation studies that will explore the predictability corresponding to specific GARP error limits in temperature, pressure, and wind.

MEMBER OF THE AUDIENCE:

How far up the line of uncertainty do you think you can go and still come down in 10 days with a 1-m/sec error?

DR. JASTROW:

We have not done the experiment, but I am almost certain that if we started at 10 m/sec, we still would come down to 1 m/sec in twice the time that was necessary for the example I discussed. There is nothing we have discovered that would indicate the contrary.

MEMBER OF THE AUDIENCE:

Would this technique work for computing winds in the tropics from satellite data?

DR. JASTROW:

Yes. I anticipated that question and have brought with me a printout of the latitudinal belts of winds.

In the tropics, the wind errors are the same as obtained here. The fraction of error is greater because the mean wind speeds are lower in the tropics, 1-m/sec error for 2- or 3-m/sec average winds. If we make use of the heavy redundancy of vertical sounding profiles and bring the temperature error down to 0.5°C, then we can determine winds everywhere to within 0.5 m/sec, and that is good enough for the tropics.

MEMBER OF THE AUDIENCE:

What is the problem in tying the cloud velocity to the wind velocity?

DR. JASTROW:

I am not well versed in that matter. I only know that some clouds do not move with the winds at all, and some move with the winds very well, as the graph shown by the previous speaker indicated. There must be an intermediate situation in which the clouds are not stationary but have some velocity relative to the winds. I do not see how you can pick that situation out.

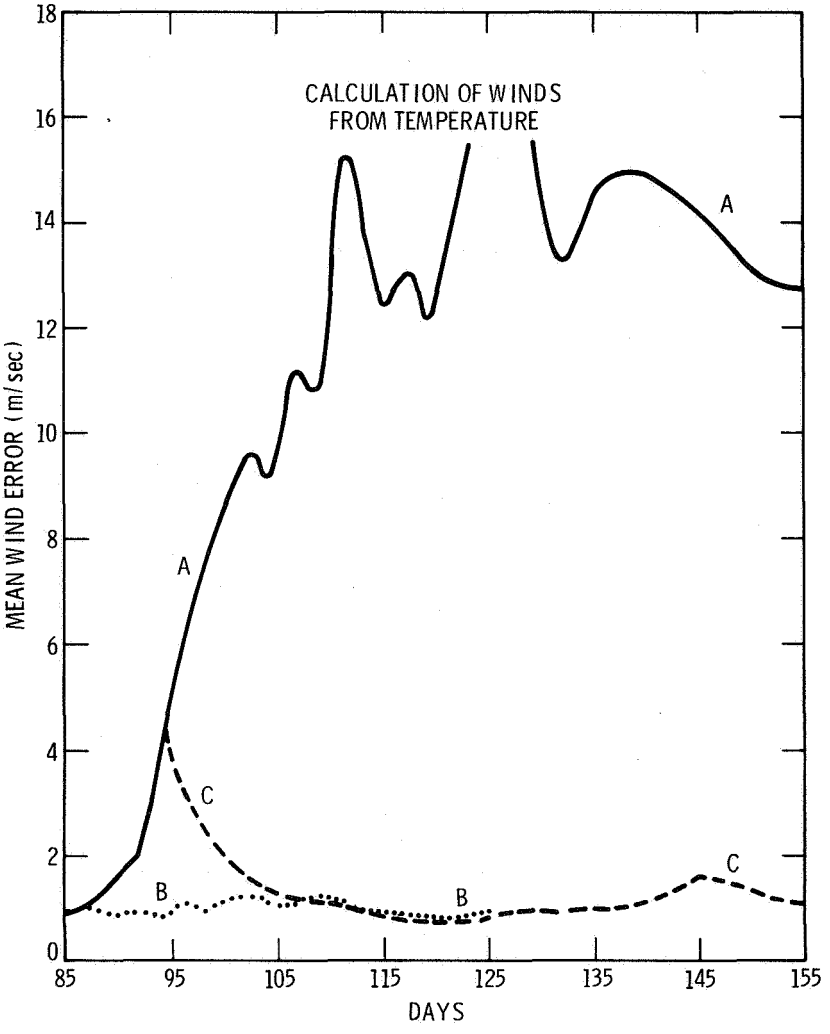

CALCULATION OF WINDS
FROM TEMPERATURE
MEAN WIND ERROR (m/sec)
DAYS
A
B
C
0
2
4
6
8
10
12
14
16
18
85
95
105
115
125
135
145
155

Figure 1

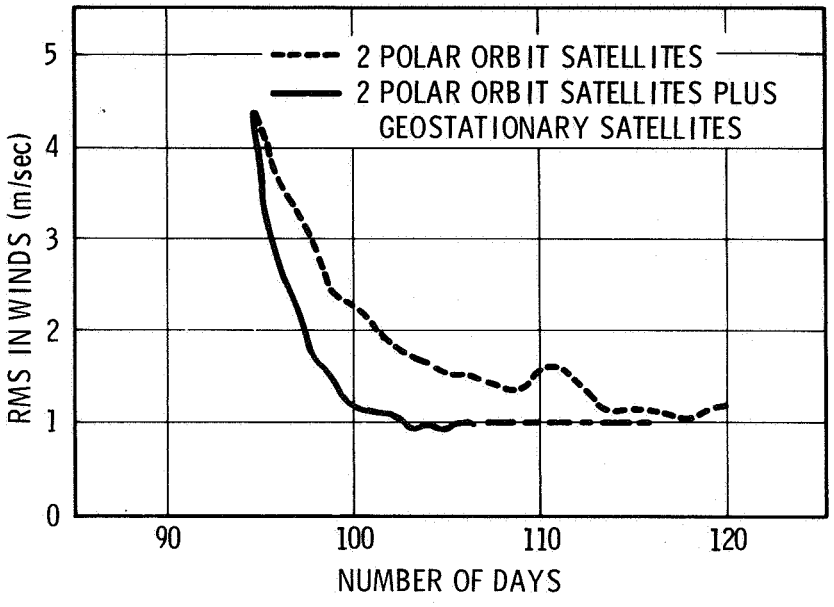


Figure 2

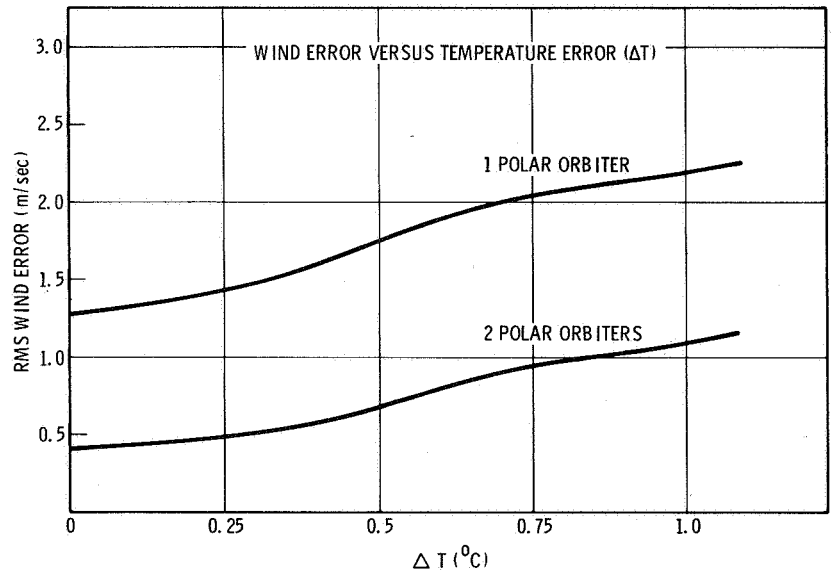


Figure 3

A NEW UV ABSORBING LAYER NEAR 50 KM

Arlin J. Krueger

CHAIRMAN:

The next paper will be presented by Mr. Arlin Krueger and concerns a new ultraviolet absorbing layer in the atmosphere.

MR. KRUEGER:

A technique used for many years to measure ozone in the atmosphere is absorption spectroscopy. Ozone is a highly effective absorber in the middle ultraviolet. The total ozone content has been measured daily at many stations by observing the attenuation of ultraviolet sunlight, and the height distribution of ozone has been determined by observing, with balloons and rockets, the increase of sunlight with altitude.

Figure 1 shows a device that we have used on ARCAS rockets. Basically, the unit consists of a filter radiometer and the associated telemetry equipment. The data are taken during parachute descent after expulsion of the device from the rocket at apogee, and the hardware, shown on the right, is used to suspend the payload. To avoid pointing problems due to the pendulation of the parachute, a transmission integrating sphere is used at the optical entrance. The ozone concentrations in a range up to 60 km change by almost three orders of magnitude. Therefore, three separate optical filters are required to make measurements in this entire range.

Figure 2 shows the results from a representative set of filters. The three active filters are at the shortest wavelengths. A filter is also provided at 3200 Å for reference purposes where the ozone absorption coefficient is very low.

We attempted to design these sets of filters such that the signals were equal at the top of the flight, which could be 50 to 55 km. We did this by adjusting the peak transmissions based on the extraterrestrial solar spectrum. In general, these attempts have been unsuccessful for filters near 3000 Å. We found that the signals below 40 km were only one-half to three-fourths what we had expected from the solar spectrum data. On the other hand, above 40 km the height gradient of intensity was much steeper than could be explained by ozone absorption alone.

Figure 3 shows the results of a sounding at Wallops Island, Virginia, in September 1968. The relative intensity data are from a filter centered at 3010 Å and a reference filter at 3228 Å.

The ozone concentrations are calculated for heights above 35 km from the two shorter wavelength filters, which are more sensitive to ozone at those heights. The curve we should have seen is shown by the dashed line. As you can see, there is a considerable discrepancy between the observed and expected intensity gradients. This discrepancy between observed intensity gradients and those calculated from ozone absorption has been observed on most of the 17 flights of this instrument. Furthermore, the same discrepancy has been found in Japanese data obtained in similar observations.

This effect appears to be limited to wavelengths near 3000 Å and probably has a bandwidth less than 200 Å. So far, we have not been able to identify the anomolous absorber, since the spectrum is unlike that of any of the well-known atmospheric constituents. The material is found in a layer sharply limited on the lower side; thus, we suspect that it may be a metastable constituent rather than a permanent atmospheric constituent.

Two excited states of oxygen are known to exist in the atmosphere at these heights: the $^1\Delta_g$ and the $^1\Sigma_g^+$ states. One of these could well have an absorption band at 3000 Å. A further possibility is vibrationally excited oxygen which should be produced at these heights from the association between ozone and atomic oxygen. To identify the absorber, the detailed absorption spectrum will have to be determined from further measurements.

The significance of this absorber is that (a) it may well be dominant at 50 km; (b) the solar flux is reduced in the stratosphere, thereby affecting equilibrium ozone concentrations; (c) the absorption may be a source of atomic oxygen; and (d) other ozone-measuring techniques in the ultraviolet may be affected.

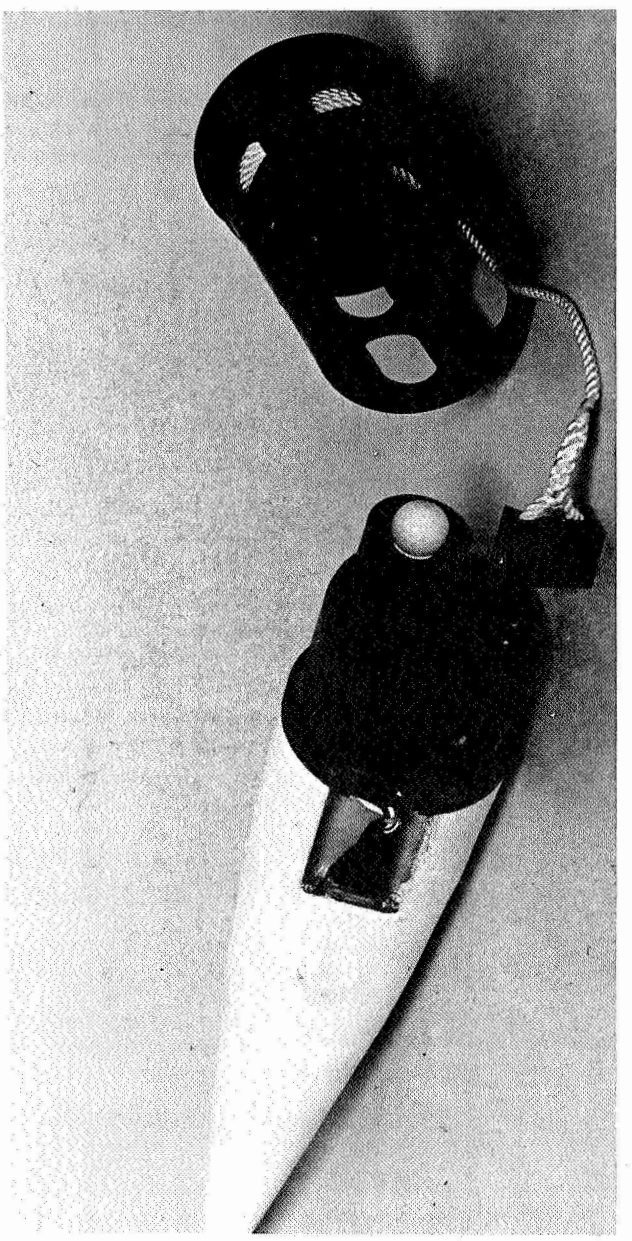


Figure 1 – ARCAS rocket ozonesonde (ROCOZ).

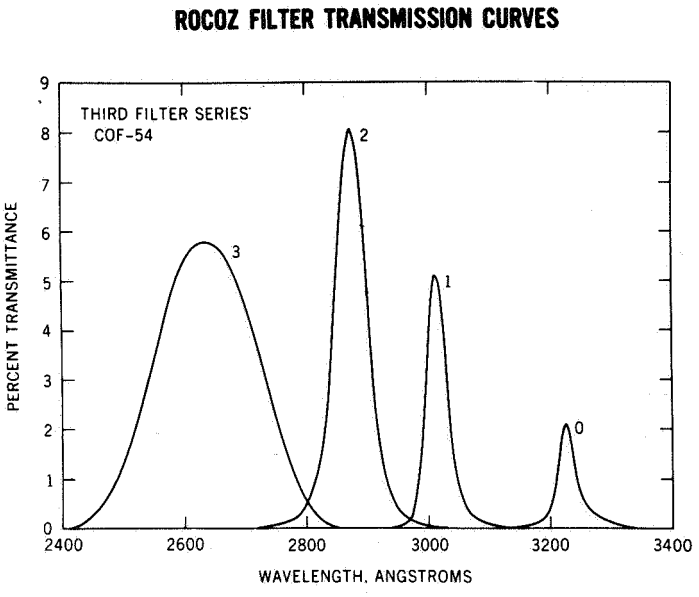


Figure 2

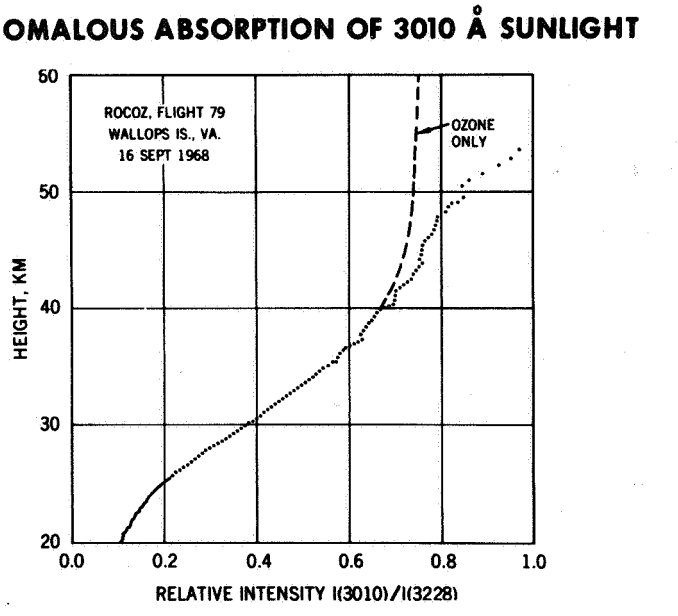


Figure 3

STRONGLY DAMPED GRAVITY WAVES IN THE EARTH'S ATMOSPHERE

Igor J. Eberstein

CHAIRMAN:

I will call next for Dr. Igor Eberstein, who will talk about strongly damped gravity waves in the earth's atmosphere.

DR. EBERSTEIN:

I will discuss the analysis of a series of vertical soundings of the atmosphere between 30 km and 120 km and its interpretation in terms of gravity wave theory.

Four Pitot tube rocket experiments were carried out over a 2-day period at Fort Churchill, Canada, in January 1967 by Smith, Theon, and Horvath. These experiments give pressure, density, and temperature with 0.5-km vertical resolution.

A mean atmosphere based on these four soundings was computed. Deviations from the mean atmosphere were calculated, and a wave-like structure was observed. The wave grows very slowly below about 80 km, then, more rapidly. Calculations of gravity wave propagation show that the experimental results may be closely approached by a plain, upward propagating, strongly damped gravity wave. It may be worth noting that an undamped gravity wave grows exponentially with altitude and the observed ones certainly do not.

Figure 1 shows the experimentally derived structure. The abscissa is altitude; the ordinate shows deviations from the mean atmosphere. You will note that the density deviation is 180 degrees out of phase with the temperature deviation and that pressure deviation is 90 degrees out of phase with density and temperature. Also, note the modulation that is observed. The wave is fairly weak and grows, becomes weak, then becomes strong again. This will be important later on. I might mention that we essentially have a snapshot of the wave in space.

Figure 2 shows calculations of gravity wave propagation in a quiescent atmosphere. You will note that the relative magnitudes of the deviations are correct and the phase relationship is correct. Density variation and temperature variation are again 180 degrees out of phase, and pressure deviations are 90 degrees out of phase.

Figure 3 shows a similar calculation but with a sinusoidal steady background wind. The wave essentially propagates into a steady wind field, and what is observed is a very strong modulation effect. It is similar to what is seen in Figure 1, only a bit more pronounced. The correct relative deviations of temperature, density, and pressure are still maintained, and the correct phase relationship is maintained.

The analytical predictions are based on a theory developed by Dr. Hans Volland at GSFC. This is an extension of the gravity wave theory due to Hines. Volland assumes that the most significant damping is due to the transfer of thermal energy. Damping by conduction, radiation, and turbulence was considered, and turbulence was found to be the most important. The relatively small-scale irregular turbulent motions provide effective mixing and heat transfer, thus smoothing out the gradients in the wave.

In conclusion: One, semiquantative agreement between an atmospheric, wave-like structure and gravity wave theory has been demonstrated; two, the most important damping mechanism is turbulence; and three, simultaneous knowledge of the thermodynamic parameters and the wind field is important.

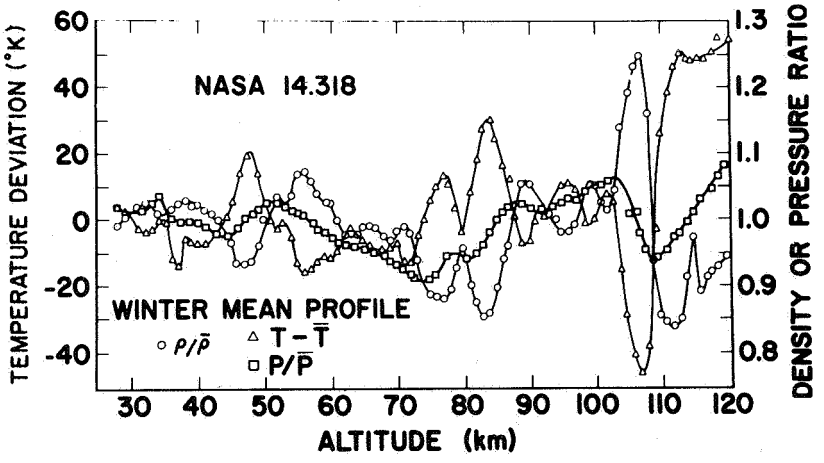


Figure 1

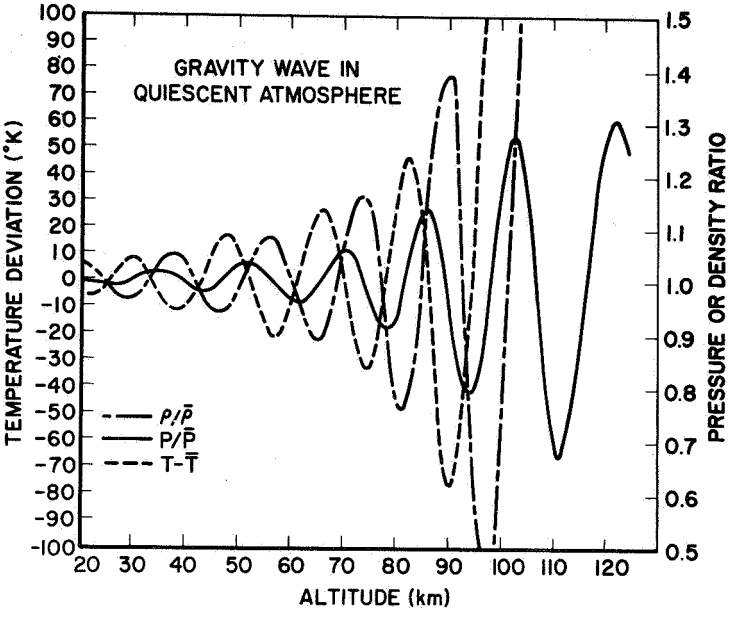


Figure 2

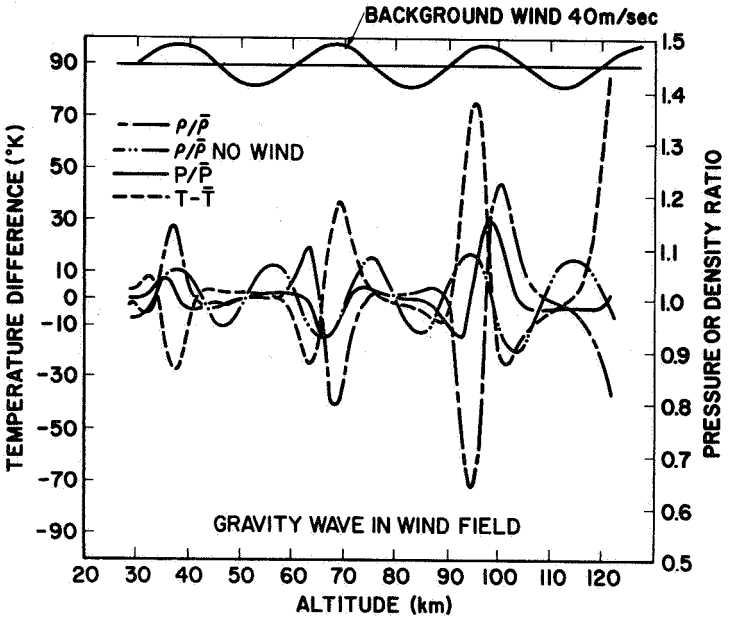


Figure 3

DIURNAL PHASE ANOMALY IN THE EARTH'S UPPER ATMOSPHERE

Sushill Chandra

CHAIRMAN:

I will call next on Dr. Sushill Chandra, who will talk about the diurnal phase anomaly in the earth's upper atmosphere.

DR. CHANDRA:

There is considerable discrepancy between the phase and the diurnal amplitude of the thermospheric temperatures inferred from satellite drag and Thomson backscatter results. The main differences in the two experiments are listed in Table I.

Table 1

Comparison of the Diurnal Characteristics of the Thermospheric Temperature

Characteristic	Satellite Drag	Thomson Backscatter
Time of maximum	1400 hours	1700 hours
Time of minimum	03-04 hours	05-06 hours
Day-to-night ratio	1.3	1.5

According to satellite drag results, the time of the maximum temperature is at 1400 hours local time. The time of the maximum inferred from backscatter results is 1700 hours local time, indicating a phase lag of about three hours. The time of the minimum inferred from the drag is from between 0300 and 0400 hours, and the time of the minimum inferred from the backscatter is between 0500 and 0600 hours, a phase difference of two hours. Again, the day-to-night ratio in temperature inferred from satellite drag is about 1.3 and the same ratio computed from backscatter is about 1.5, so even this is not in agreement.

Because of this discrepancy some serious questions are now being raised about the validity of inferring the true diurnal variations from satellite drag measurements. This is particularly so because both the theoretical calculations and the direct measurements of the neutral temperatures from OGO 6 show results more in accord with the backscatter data.

The purpose of this presentation is to show that observational data do not contradict each other if it is recognized that the basic parameter measured from the drag experiment is neutral density. The temperature is only a derived quantity. In the case of backscatter, the temperature is the basic parameter. The problem, therefore, is in computing the temperature from the density, and vice versa. It is proposed here that the concept of dynamic diffusion, as opposed to the static diffusion usually employed in computing the temperature from the density, leads to a reasonable reconciliation between the two sets of observational results.

The basic concept of dynamic diffusion involves the assumption that the neutral constituents of the atmosphere are in diffusive equilibrium, even though the temperature and the density of the lower boundary are not invariant. The changes in the lower boundary are caused by thermally excited winds in the lower atmosphere, tidal forces, and eddy diffusion.

Figure 1 shows the diurnal variations in density and temperature at 300 km, as obtained by solving the heat conduction equation and using the concept of dynamic diffusion. The density of O_2, N_2, and O at 100 km is assumed to be of the form $1 + 0.3 \cos \omega t$, where ω is the angular frequency of the earth and t is the time measured from noon.

Thus the maximum value will be at noon, when it is $1 + 0.3$, or 1.3 times the mean value, and the minimum value of the density will be at midnight, when it will be 0.7 times the mean value.

Thus the day-to-night changes at 100 km are assumed to be a factor of 2.

The result of this calculation shows that the temperature, which is the dashed line in Figure 2, has its maximum value at 1700 hours local time. The corresponding density maximum at 300 km occurs at around 1400 hours local time.

Again, the temperature minimum occurs at about 0500 hours and the density minimum occurs at about 0300 hours. Both are consistent with the backscatter and drag results.

The day-to-night variation in temperature is from about 800°K minimum temperature to about 1200°K maximum, which is a factor of 1.5, and the density changes are found to be about 2.0×10^{-14} to 6.0×10^{-14} gm cm^{-3}, which is a factor of 3 increase. Thus, this is also in good agreement with the satellite drag results.

We have carried out extensive calculations between 100 and 1500 km on a global basis, for both density and temperature, and we have found that the concept of dynamic diffusion can be usefully employed to satisfy both sets of observational data.

In conclusion we therefore note that the introduction of the dynamic diffusion concept has removed the basic discrepancy in satellite drag and backscatter results. It is suggested that in computing the temperature from the density, the assumption of the invariant boundary is not correct. Since this is the basic assumption in constructing the international reference atmosphere, it is suggested that both the international reference atmosphere and the U. S. standard atmosphere should be thoroughly revised to make them compatible with observational data.

CHAIRMAN:

Are there questions for Dr. Chandra?

MEMBER OF THE AUDIENCE:

This paper illustrates the importance in making actual measurements near the 100-km level which hopefully will be made by atmospheric explorers.

DR. CHANDRA:

Yes.

MEMBER OF THE AUDIENCE:

Could that then be used to get a boundary condition for the wind system?

DR. CHANDRA:

Yes, exactly. The atmospheric explorer hopefully will give us some indication on that line, yes.

DIURNAL VARIATIONS IN THE NEUTRAL ATMOSPHERE AT 300 Km

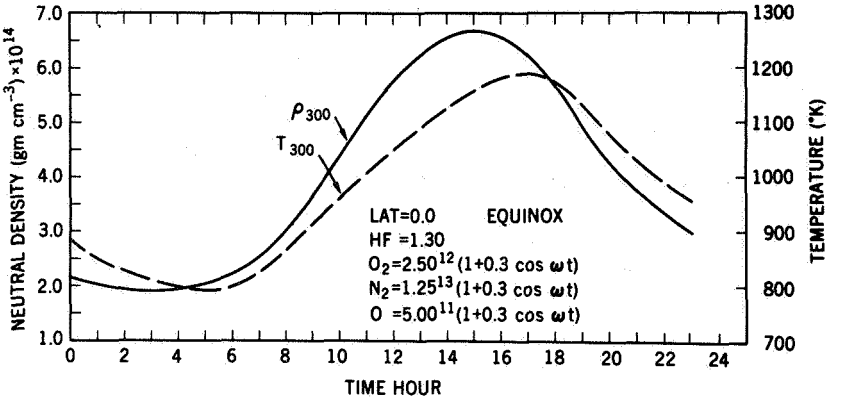


Figure 1

THE EFFECT OF ATMOSPHERIC WINDS ON THE O^+-H^+ TRANSITION LEVEL

Henry C. Brinton

CHAIRMAN:

Next I will ask Mr. Henry Brinton to tell us about his work on the effect of atmospheric winds on the O^+-H^+ transition level.

MR. BRINTON:

For some time it has been thought that neutral winds at F-region altitudes, caused by pressure gradients in the neutral atmosphere, may have a significant effect on the ionosphere. I am going to present data describing the global variation of the O^+-H^+ transition level, that is, the altitude at which the dominant ionospheric constituent changes from O^+ to H^+. We interpret certain features of this variation as further evidence that winds in the neutral atmosphere, through neutral-ion drag, may, in fact, strongly affect the ionosphere.

In Figure 1 we have plotted the altitude variation of the O^+-H^+ transition level as a function of geomagnetic latitude at 3 longitudes. These observations are from the ion mass spectrometer on the Explorer 32 satellite, obtained during the summer of 1966 at midday. We note that at European and Pacific longitudes there is approximate symmetry about the geomagnetic equator, whereas at American longitudes the transition level is higher by several hundred kilometers in the Northern Hemisphere than in the Southern Hemisphere. It is this symmetry at European and Pacific longitudes and asymmetry at American longitudes that we attribute to the effects of neutral winds.

In Figure 2 we show at the top the magnetic dip angle configuration on a global scale. Note the similarity between the dip angle configuration at Pacific and European longitudes and the downward shift of the dip angle contours at American longitudes. You will recall that the observations shown in Figure 1 were obtained during Northern Hemisphere's summer, when the sun was at an average geographic latitude of +13°, shown here by the dashed line. This means that the sun was symmetric in the magnetic field reference frame at Pacific and European longitudes, but the subsolar point occurred in the Northern Hemisphere at American longitudes at a latitude of about 20° north.

It is this asymmetry of the sun in the magnetic field reference frame at American longitudes and the resultant asymmetry of the neutral winds that we invoke as the mechanism causing the asymmetry in the observed transition level.

Following the earlier work of Kohl and King, Volland and Mayr at Goddard have derived the neutral wind field from the Jacchia neutral atmosphere model, and we show certain features of that derivation at the bottom of Figure 2.

At European and Pacific longitudes, the subsolar point lies on the magnetic equator for our observations. The horizontal component of the neutral wind increases with increasing distance from the subsolar point north and south, and we see that there is symmetry of this horizontal wind component in the northern and southern magnetic hemispheres at these longitudes. At American longitudes, however, the subsolar point lies in the northern magnetic hemisphere, and the same wind distribution about the subsolar point causes an asymmetry in the horizontal wind component in the magnetic reference frame, with a larger horizontal component observed in the Southern Hemisphere than in the Northern at the same latitude.

The downward component of the wind parallel to the magnetic field lines can account for the observed asymmetry in the transition level. To demonstrate how this model works, Figure 3 shows the results of the calculation of the hydrogen and oxygen ion concentration profiles at American longitudes.

These profiles were derived by solving the ion and electron continuity equations along a field line intersecting the equator at 2500 km altitude, taking the wind field into account. In the Southern Hemisphere, where the downward wind component is larger, the O^+ ions are pulled downward by neutral-ion drag to lower altitudes where the O^+ can recombine. This loss of O^+ causes an ion sink to be created in the Southern Hemisphere, which then draws hydrogen ions up along the field line in the Northern Hemisphere and down along the field line in the Southern Hemisphere. The result of this proton flux is an increase in the transition level in the Northern Hemisphere and a decrease in the Southern.

The calculation, then, is in agreement with our observations, in that the transition level is several hundred kilometers lower in the south than in the north at this longitude.

MEMBER OF THE AUDIENCE:

How much change would you get in this transition level in a purely diffusive model as a result of a difference in temperature at the subsolar points if they are quite some distance away?

MR. BRINTON:

I am not clear on the question. Considering the asymmetry in the field?

MEMBER OF THE AUDIENCE:

Yes.

MR. BRINTON:

I do not know. I would have to look at that.

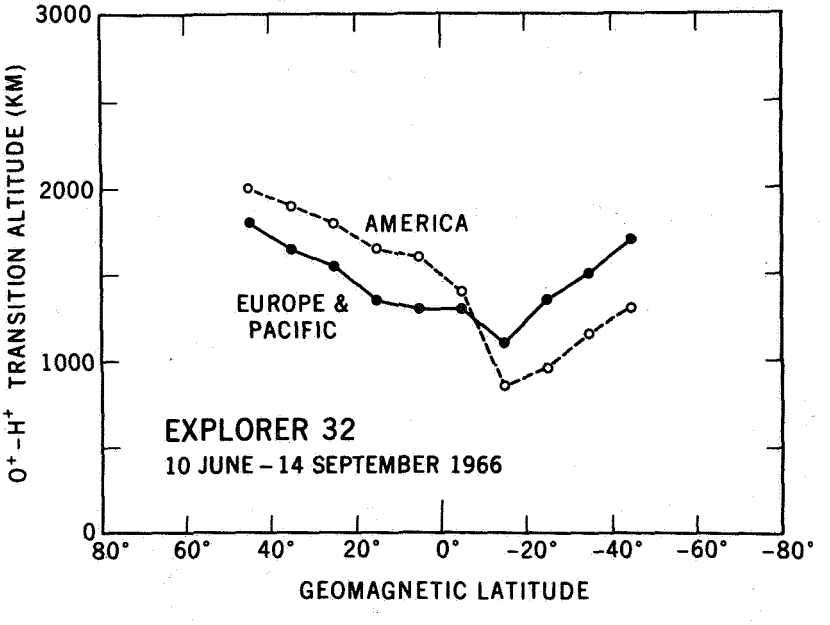


Figure 1

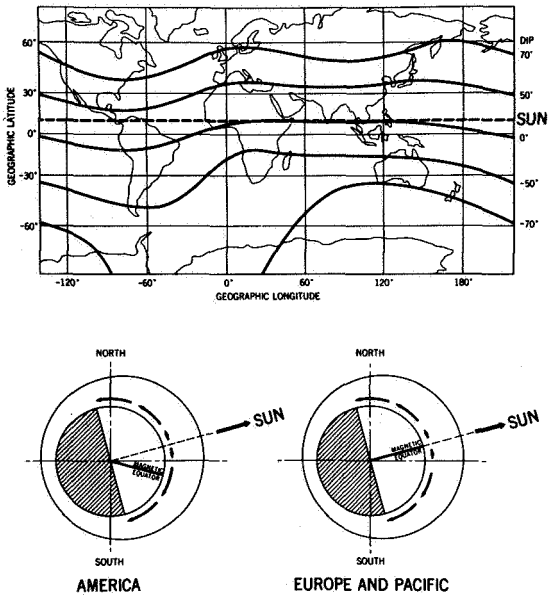


Figure 2

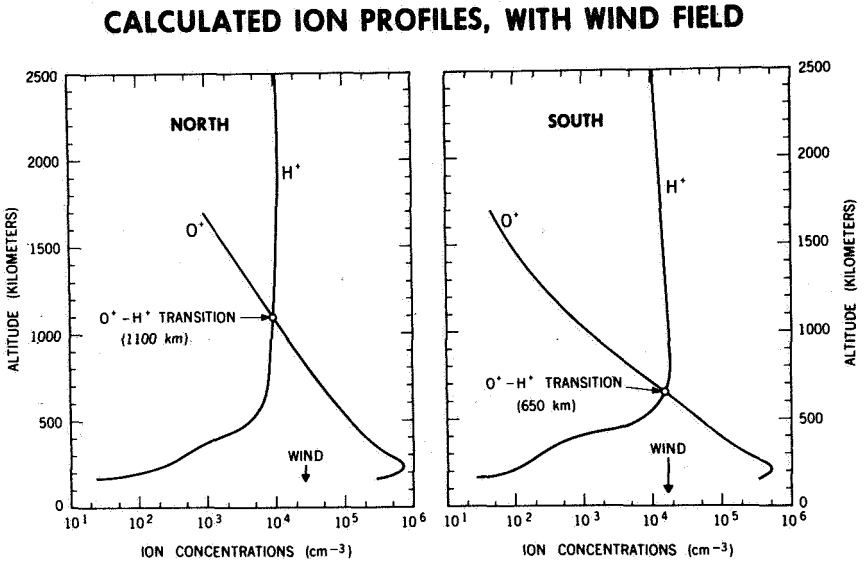


Figure 3

LATITUDINAL DISPERSION OF GASES IN THE UPPER ATMOSPHERE

Carl A. Reber

CHAIRMAN:

We will continue our discussion of latitudinal dispersion of gases by calling next upon Mr. Carl Reber.

MR. REBER:

The horizontal distribution of gases in the upper atmosphere has been a topic of some interest since the discovery of the so-called winter helium bulge. The winter helium bulge refers to the fact that the helium abundance in the upper atmosphere is an order of magnitude higher in the winter hemisphere than it is in the summer hemisphere, and this was one of the topics studied in some detail with the mass spectrometer on the OGO 6 satellite.

Now, before we get into the anomolies in the atmosphere, it might be interesting to look at what we would expect to see if we had a spherically symmetric atmosphere.

In Figure 1, we have the satellite trajectory through our symmetrical atmosphere, and we show, too, constant density or pressure surfaces C_1 and C_2. Because different constituents have different scale heights, these constant pressure surfaces will be closer together for heavier constituents like N_2 and O than they will be for lighter constituents such as He or H. The relative altitude behavior is shown in Figure 2; the He has a large scale height, and the N_2 has a rather low scale height.

As the satellite passes through this symmetrical atmosphere, we would expect to see a signal generated such as that shown in Figure 3. When the satellite moves in toward perigee from the higher altitudes, the signal increases above the background level, and at perigee, we observe the maximum density of all three gases which then decreases toward the background level once more. The point to be made is that the symmetry of the atmosphere produces the peak for all three gases at the same location, the perigee.

Figure 4 shows what we would expect to see in a nonsymmetrical atmosphere. In this case we have drawn the N_2, which represents the major portion of the atmosphere, with its peak in the Northern Hemisphere and the

He with its peak in the Southern Hemisphere. *P* indicates the perigee of the satellite orbit. As the satellite passes into the lower portion of the atmosphere, we first would observe the maximum N_2 signal fairly close to the perigee and close to the actual N_2 peak itself. Further through the orbit we would observe the peak in the He density measurement at a point which lies roughly halfway between the perigee and the peak of the actual He density.

Figure 5 shows data obtained from OGO 6. It can be seen that, in fact, there is a separation of the densities of the major gases from the He density, with the He peak occurring later in the orbit. The perigee was between the N_2 and the O peaks, and the difference in latitude between the peaks of the major gases and the He was about 50 degrees.

The main point is that now we are not only refining our information and our understanding of the distribution of He in the upper atmosphere, but, in addition, we are observing a latitudinal separation for the two major heavy components, O and N_2, which is different from that of He.

CHAIRMAN:

There appears to be a difference in the distribution of O and N_2 heretofore unsuspected. These are OGO 6 data which were just recently obtained.

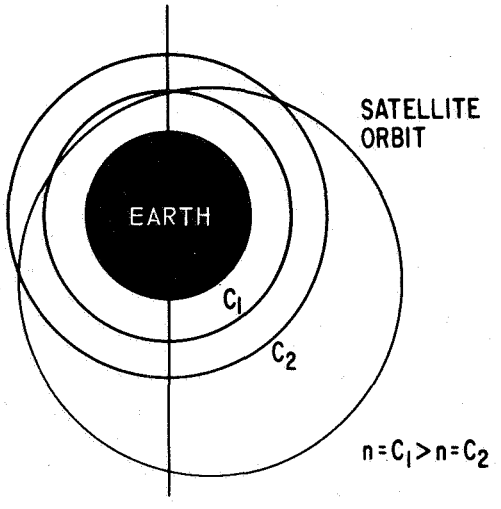


Figure 1

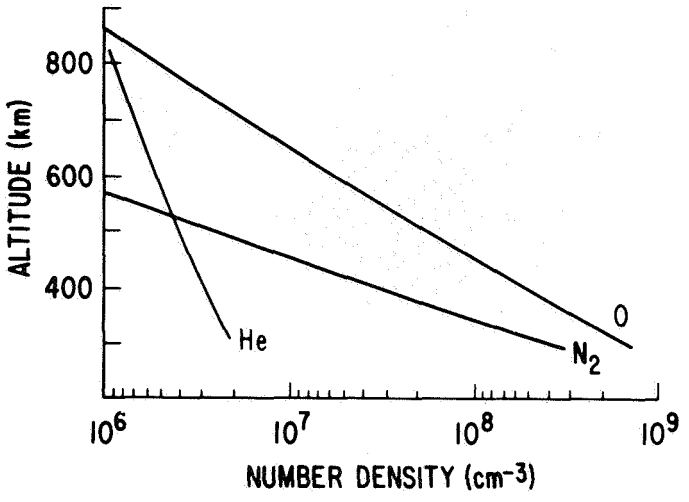


Figure 2

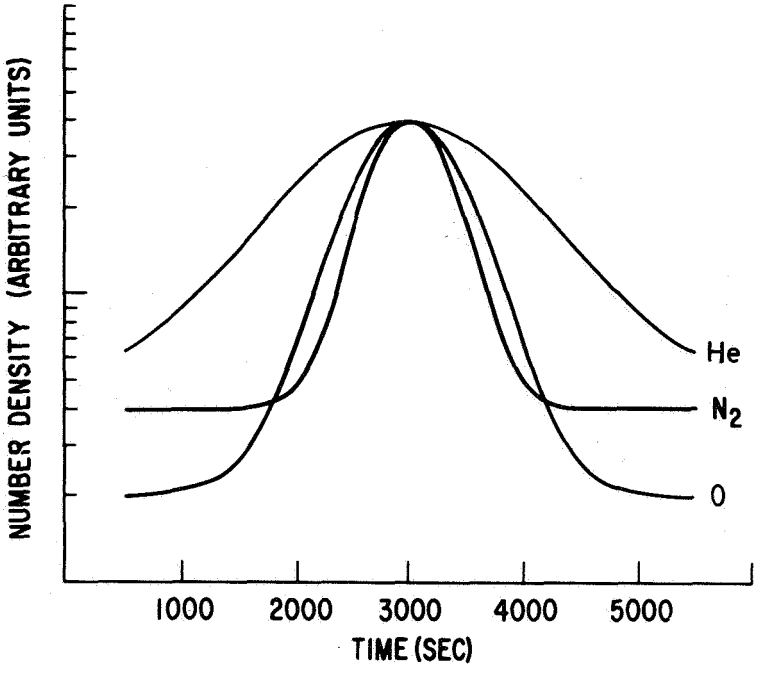


Figure 3

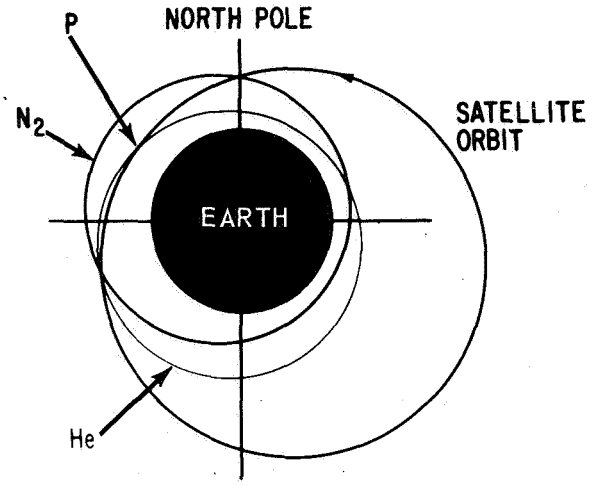


Figure 4

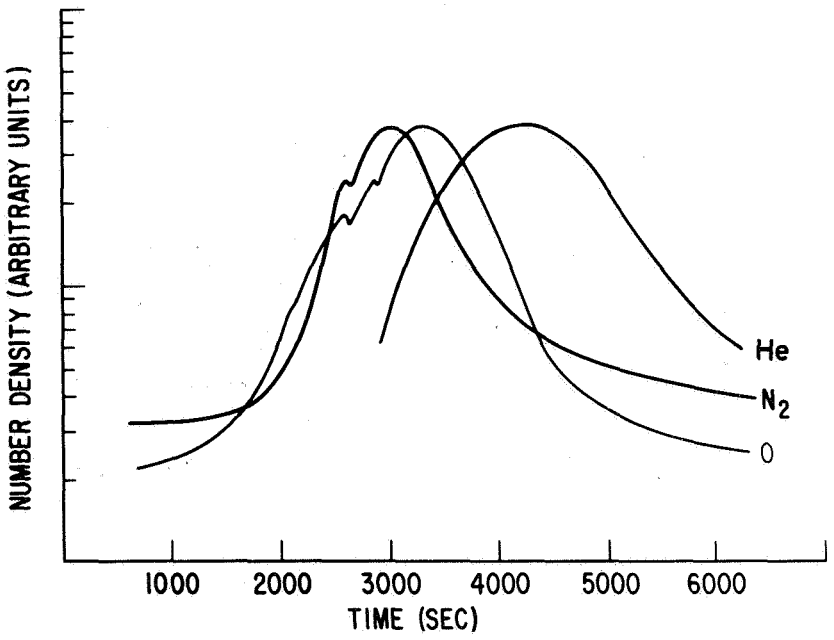


Figure 5

AIRGLOW OBSERVATIONS FROM OGO 4

Edith I. Reed

CHAIRMAN:

Next, we will have a discussion of airglow observations from OGO 4, presented by Mrs. Edith Reed.

MRS. REED:

Using the airglow photometers on OGO 4 we have observed several principal airglow and auroral emissions. Today we will talk about just one, the 6300-Å red line of atomic oxygen, and we will talk only about the phenomena at less than 40 degrees of latitude. We will look at a map of the airglow in the equatorial region and indicate how this is related to the neutral and the charged components of the atmosphere.

First let us briefly review the processes that are thought to be responsible for the formation of the red airglow.

First we have a charge exchange reaction

$$O^+ + O_2 \rightarrow O_2^+ + O$$

in which molecular oxygen ions are formed as a result of charge exchange with atomic oxygen ions. This is followed by a dissociative recombination, resulting in two oxygen atoms, one of which may be excited:

$$O_2^+ + e \rightarrow O + O^*{}_{6300\ \text{Å}}$$

This excited atom, after about 110 seconds, will emit a 6300-Å photon. However, if during this time there are a sufficient number of collisions with nitrogen molecules, the excited atom loses its energy through collisional deactivation and does not emit a photon. Hence, this 6300-Å emission is confined principally to the F-region.

Above about 250 km we always have some red airglow. This red airglow should be proportional to the number of atomic oxygen ions, the number of electrons, and the number of oxygen molecules that are available. It is diminished by the number of collisions with nitrogen molecules, resulting in the removal of the energy.

On OGO 4 we had a photometer that looked straight down at the surface of the earth above practically all of the airglow emission, so that at night we could measure all of the red airglow that was emitted in the atmosphere.

In Figure 1 we have combined the data for about three days in late October. These data were taken at a local time of about 9:40 in the evening.

We can see that we have an arc of the red airglow around the world with intensities up to more than 1,000 rayleighs. In the Southern Hemisphere, there is at least part of an arc that is also roughly parallel to the geomagnetic equator. It happens to be rather weak across the Pacific, with the intensity dropping to less than 50 rayleighs.

These arcs are very closely related to the ionospheric equatorial anomaly which is the characteristic of the daytime F-region. This is expected since the airglow is proportional to the square of the electron density.

However, one quickly notices that over much of the Pacific the southern arc is almost absent. To study this further, we looked at the atomic ion densities which were measured on the same satellite with an ion trap experiment.

Dr. S. Chandra and B. Troy of GSFC have made available to us their data for a number of passes corresponding to the airglow data in Figure 1. Let us take a look at a pass at a longitude of 90 degrees west. A comparison of the data is shown in Figure 2.

The ion densities were measured at an altitude between about 418 and 450 km. This is close enough to the F-region where the airglow is emitted so as to be representative of ion densities at the somewhat lower altitudes.

To the north we see a maximum in the ion density and a maximum in the red airglow, both at about the same L-shell. This is representative of most of the comparisons that we have made. But, perhaps 20 percent of the time, as we see in the south, there is a maximum in atomic ion density and hardly any corresponding maximum in the red airglow.

This particular region must be influenced not only by the charged particles but also by the neutral particles. It may be that the composition of the atmosphere in the Southern Hemisphere is different; that there is an insufficient number of oxygen molecules available for the charge exchange reaction. It may be that the density has increased such that there is a greater number of collisions with nitrogen molecules. Then, the energy would be

lost through collisional deactivation rather than through the emission of a photon.

So, in any case, we have another key that needs to be used in the study of the dynamics of the neutral and the charged components of the atmosphere.

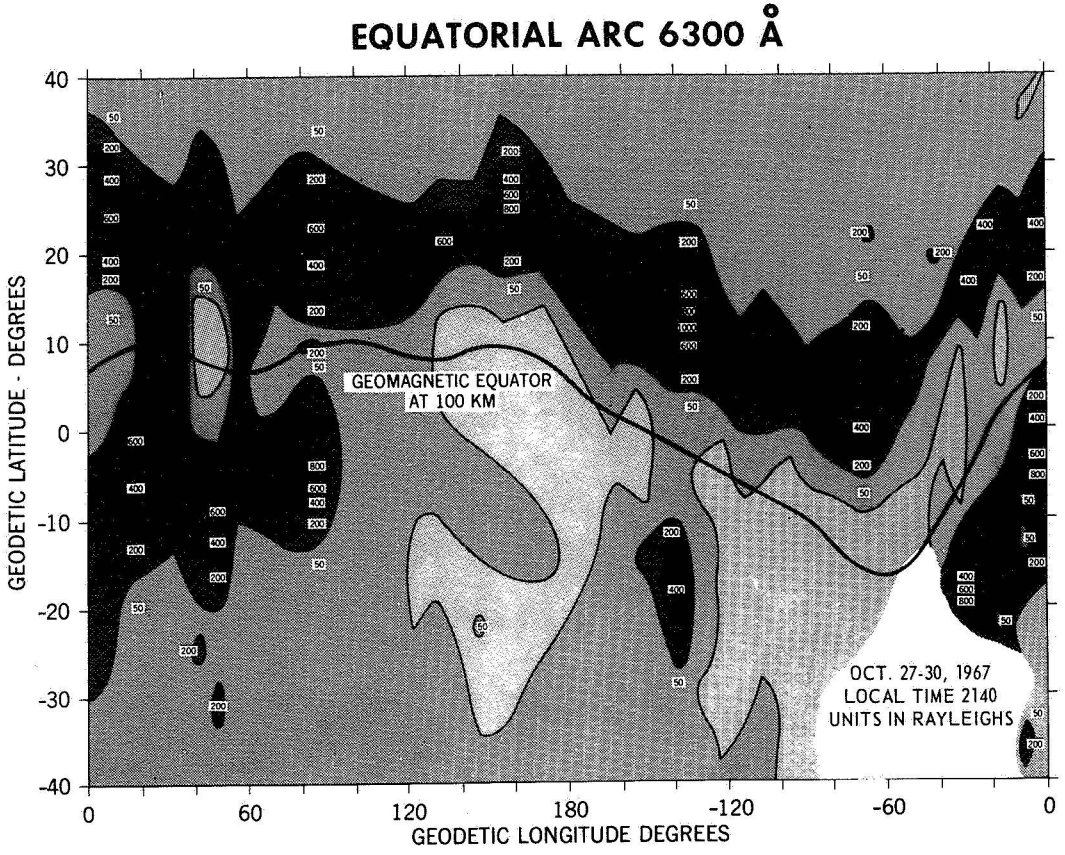

EQUATORIAL ARC 6300 Å
GEOMAGNETIC EQUATOR
AT 100 KM
OCT. 27-30, 1967
LOCAL TIME 2140
UNITS IN RAYLEIGHS
GEODETIC LATITUDE - DEGREES
GEODETIC LONGITUDE DEGREES
40
30
20
10
0
-10
-20
-30
-40
0
60
120
180
-120
-60
0

Figure 1

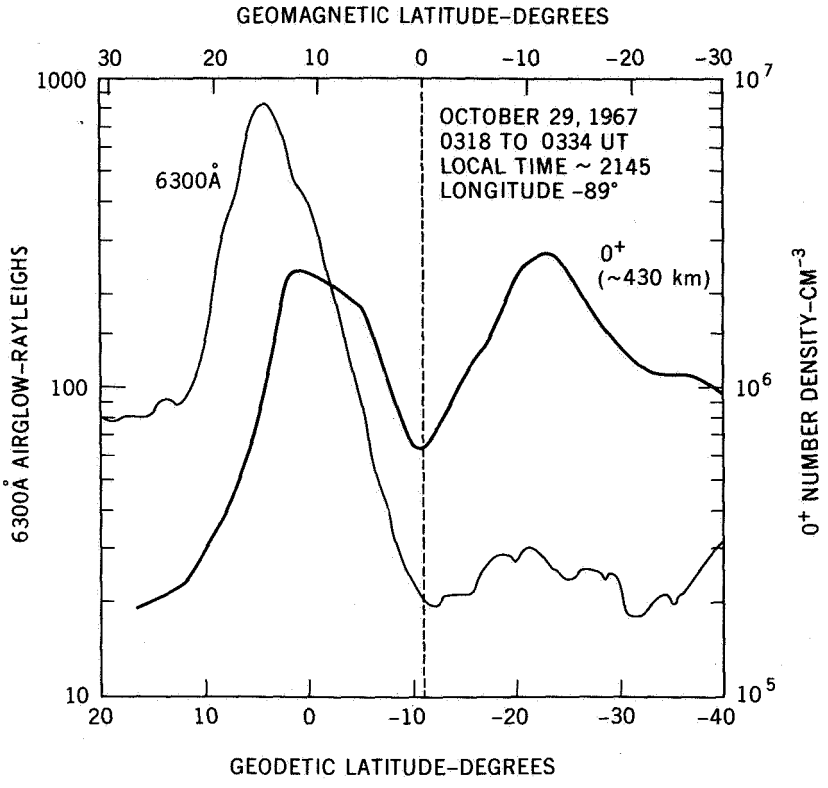


Figure 2 – OGO 4 airglow and ion trap data.

THE ARTIFICIAL AURORA EXPERIMENT

John C. Price

CHAIRMAN:

To continue the session, we will call for the paper by Dr. John Price on the artificial aurora experiment.

DR. PRICE:

The artificial aurora experiment was based on a rocket flight from Wallops Island, Virginia, on the night of January 26, 1969. The experimental situation is shown in Figure 1.

The rocket was flown to an altitude of approximately 250 km. It carried on it an electron accelerator which was to fire a programmed sequence of pulses downward along the local magnetic field line. The idea was that the electron beam, when it hit the Earth's upper atmosphere, would cause an auroral spot that would be easily visible with photographic equipment and possibly to the naked eye.

This is a very interesting experiment in terms of space physics as well as plasma physics, because it has been known for some time that a configuration like this is unstable with respect to electrostatic instabilities.

Figure 2 shows some of the properties of the two-stream instability, which is the instability that would be expected to be present in this case.

Here we have the distribution of electrons, the number of electrons at a given velocity in a region where the beam is passing through. You can see that there is a great number of background electrons which are essentially cold; they have very low energies. Passing through the volume element is a very-high-energy 10-kV beam which is going downward toward the ground.

In this case, plasma theory would predict that any electric field that could be expected to be present in the system initially as a fluctuation would grow exponentially, with the growth rate given in terms of the beam current, the area of the beam, the velocity of the beam, and the density of of the background plasma.

This is an interesting problem because we do not really know what the answer is in an exact sense. If the electric field grows exponentially, we

can expect after a short time—this turns out to be quite a rapid growth rate—that the strong electric field that will be present will scatter the beam. The beam would no longer follow the field line down to the Earth's atmosphere, and no spot would be visible.

In the experiment, the accelerator emitted a sequence of pulses that varied in strength and duration. Four of the strongest spots were observed, as reported by Dr. Hess earlier this year in *Science* magazine.

In the naive sense, the beam would not go down. I would like to propose as a model that the beam, in the process of thermalizing or being scattered by this background gas of cold electrons, gives its momentum to the background electrons and hence causes a snow plow, or slug of slightly heated electrons, to sweep down the field line picking up more background electrons as it goes. Once the snow plow has cleaned out the background electrons, the spot should be visible, at least to the photographic equipment.

In Figure 3 we see that the situation is not quite so simple because in addition to the background electrons there are also background neutrals, as shown in the right column.

The neutral density falls off very abruptly at higher altitudes. The plasma density increases from 10^4 electrons/cm^3 (it is nearly zero below 100 km); it then goes up quite rapidly toward 10^5 electrons/cm^3 and beyond at higher altitudes. If one calculates the difference between the growth rate (from the plasma instability) and the damping rate (from the neutral density), it is seen that the beam is unstable only in the higher ranges, quite near the rocket. From there on down it is stable.

Thus, if you use the model I have just proposed, which is the snow plow pushing the background electrons out of the way, you can make some estimates as to when the beam should be visible. In this case you find that only the strongest pulses, the 10-kV pulses, should be visible. The strongest pulses, which were signature pulses, lasted one second, and these were recorded by the photographic equipment. None of the lower-energy pulses were observed. Some of the 0.1-second pulses possibly were observed; it is not quite clear.

The rocket did not fly the expected trajectory, so the photographic equipment was not looking at exactly the correct place. However, if the snow plow model is correct, there should be a time delay before the pulses are visible. According to a preliminary report, there should have been a 60-msec time delay. I have been told since then that it is now consistent

with no time delay for the photographic results. The model indicates that a time delay on the order of 20 or 30 msec would be expected. However, there is so much energy loss due to the instability that the auroral spot would be just barely visible in any event, and this apparently agrees with the photographic results.

Finally, if we look at this same type of configuration and consider how far this given accelerator could go up the field line and shoot a pulse down, we would estimate that, at an altitude of 400 to 500 km, the snow plow would have so much background density to sweep out the pulse would be marginally visible. A longer path length would result in no visible spot at all.

MEMBER OF THE AUDIENCE:

Was there any manifestation of radiofrequency emission associated with experiment?

DR. PRICE:

No. I do not know of any radiofrequency results. There was a collector that was sensitive to rather high-energy electrons, that is, those electrons in the background plasma heated by the beam. These electrons were observed, and the heating was significant when compared with that from collisions with neutrals.

MEMBER OF THE AUDIENCE:

Does the beam energy go into heating and not radiofrequency?

DR. PRICE:

Well, we can say it goes into heating certainly. The radiofrequency was not monitored, to my knowledge.

DR. BAUER:

There was an experiment to be performed in conjunction with the electron beam. A second rocket was instrumented to measure the radiofrequency emitted by the beam interaction. Unfortunately, the rocket made measurements only to about 20 km altitude. There was, I understand, low frequency observation equipment on the ground, but I am not aware of any results from it.

DR. CLARK:

Could you be a bit more specific about the snow-plow mechanism? Is that simply scattering of the thermal electrons out of the beam?

DR. PRICE:

The strong electric fields generated by the instability slow down the beam and very slightly accelerate the background. Since this happens very very rapidly, it would be expected only at the front of the electron beam. So, as the beam comes down, there is a slug created at the front when the beam strikes from above, and it pushes the cold background as it goes down.

DR. CLARK:

So this all goes to show that not all experiments work the way you want them to work the first time. It also, I think, indicates that, before one were to try the experiment that Dr. Hess originally intended, of shooting the beam from one hemisphere to the next, it would be desirable to try it again from a higher altitude.

DR. PRICE:

I would definitely suggest that.

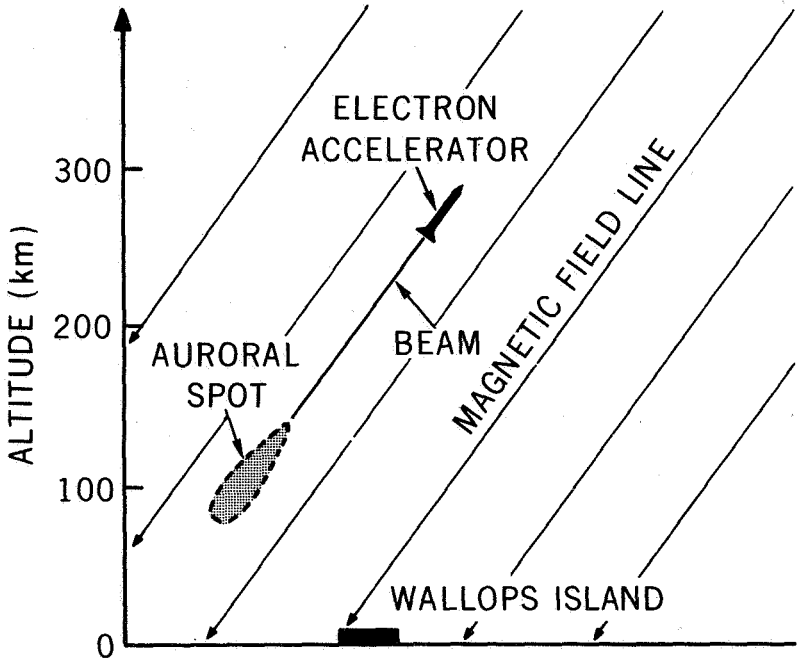


Figure 1

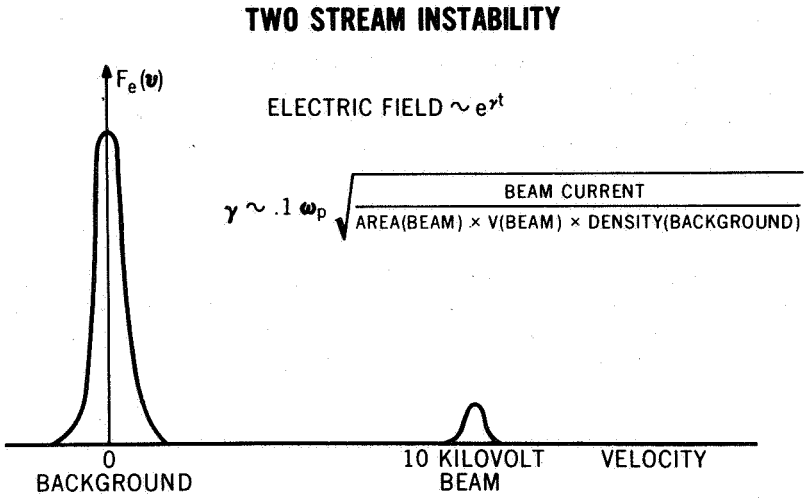


Figure 2

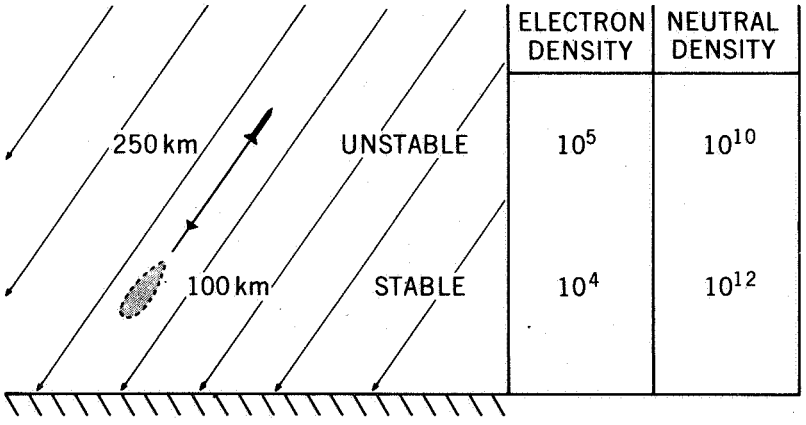


Figure 3

THERMAL PLASMA NEAR THE PLASMAPAUSE

Joseph M. Grebowsky

CHAIRMAN:

Next, Dr. Joseph Grebowsky will tell us about thermal plasma near the plasmapause.

DR. GREBOWSKY:

At a geocentric distance of a few earth radii near the equatorial plane, a rapid drop of the ambient electron density usually occurs with increasing altitude. This region of change is commonly referred to as the knee or the plasmapause. Across the same group of magnetic field lines at altitudes near 1000 km, a change in plasma density is also observed in the form of troughs in the latitudinal distribution of the light ions and electrons. I intend to show that associated with these density variations is a change in the plasma's state of motion along the field lines.

Figure 1 depicts a schematic model of the mechanism responsible for the formation of the plasmapause. Electric field drift of plasma out of closed magnetospheric field lines results in a loss of plasma from the outer magnetic flux tubes. Such a loss mechanism does not exist in the inner region, the plasmasphere, which consists essentially of trapped plasma that co-rotates with the earth. Hence the densities in the outer flux tubes can be much lower than the plasmasphere density. Typically, at the equator the density is of the order of 5 particles per cm^3 on these outer closed field lines. Also implicit in this model is the existence of fluxes of plasma which continually move upward on the outer field lines as ionospherically produced plasma replenishes the plasma loss.

Within the plasmasphere, however, no loss occurs via transport out of the magnetosphere; if fluxes do exist along the field lines, they arise from diurnal effects and correspond to proton flow velocities which are small in magnitude and which can be directed either away from or toward the earth. By contrast, the proton flow velocities in the outer flux tubes may attain supersonic values. To test the validity of this model, model computations were made and the results compared with experimental measurements. In Figure 2, a trough in the H^+ ions as measured by OGO 4 at 900 km just before dawn is shown. In contrast to the measured H^+ decrease, the measured O^+ density increased with increasing latitude.

Starting from a base-level altitude in the topside ionosphere at which latitudinal variations could be ignored, the plasma continuity and the plasma momentum equations were integrated along the field lines to compute the densities at 900 km. If it is assumed that no plasma motion exists anywhere along the field lines, the computed densities—indicated by the heavy dashes—show very little latitudinal variation, which is in sharp contrast to the measurements. However, we can reproduce the measured densities if it is assumed that a downward flow of plasma exists along low latitude field lines corresponding to the collapse of the outer plasmasphere at night, and, furthermore, that an upward-directed flux exists along the high latitude closed field lines in accordance with the previously discussed model.

Figure 3 shows that, if we extrapolate these results to the equator, we must assume the existence of particle motions along the field lines in order to produce the characteristic low densities of 5 particles per cm^3 or less that are known to exist at the equator on the outer field lines. A stationary state of plasma everywhere cannot produce these low densities.

Hence, in conclusion, it appears that the termination of the plasmasphere is associated with a change in the plasma state of motion along the field lines, with continuously upward-directed fluxes existing on the outer closed field lines.

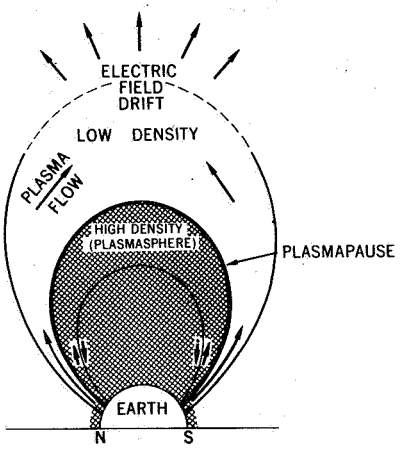


Figure 1

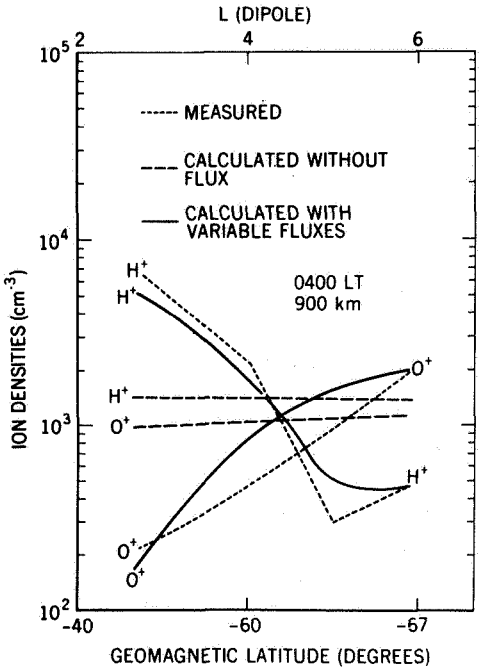


Figure 2

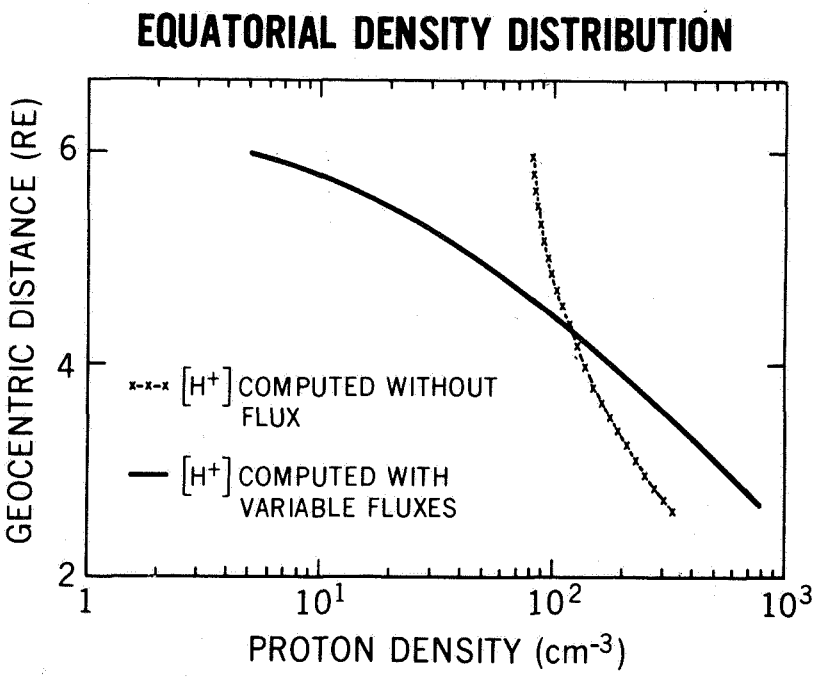


Figure 3

THE INFLATION OF THE INNER MAGNETOSPHERE

Masahisa Sugiura

CHAIRMAN:

The next paper concerns the inflation of the inner magnetosphere and will be presented by Dr. Masahisa Sugiura.

DR. SUGIURA:

Studies of the distortion of the magnetic field in the inner magnetosphere have been made, in the past, mostly relative to magnetic storms. This is because observational data with which the inner magnetospheric field can be studied accurately have not been available.

Now the OGO 3 and OGO 5 satellites are providing extensive and accurate data in this area as well as in regions at greater distances from the Earth. The results presented here are based on sample data from these satellites; these results are continuously being updated and expanded as the analysis progresses with time.

For a description of the distortion of the magnetospheric field, quantity ΔB, defined as the observed minus the reference field magnitude, provides a useful measure. Here, the reference field is based on the spherical harmonic expansion of the main field presented by Cain and others.

Figure 1 is an example of contour maps of the ΔB distribution, this one representing ΔB distribution in the geomagnetic noon-midnight meridian plane for a slightly disturbed condition, with Kp equal to 2 and 3.

This contour map and similar maps for different degrees of disturbance and different local times have shown the following two important features: the first is the existence of a region of large positive ΔB at middle to high latitudes; the second is the large decrease of the field in the inner magnetosphere. These features are observed even during magnetically quiet or only slightly disturbed periods.

To illustrate the existence of this positive ΔB region, the ΔB values taken in the nightside magnetosphere between 8 and 12 Earth radii in geocentric distance and within 4.5 hours of geomagnetic midnight are plotted in Figure 2 as a function of geomagnetic dipole latitude. High ΔB values are clearly indicated above geomagnetic latitude 20° or 25°. Different symbols are used to represent different degrees of disturbance. Those points

obtained during magnetically quiet periods are shown by dark dots, those obtained during slightly disturbed periods by open circles, and those during magnetic storms by triangles. It is evident that ΔB increases progressively with increasing degree of disturbance.

The average of the ratio of ΔB to B_0, where B_0 is the magnitude of the undisturbed reference field, for the points with $\Delta B \geq 30\ \gamma$, is 0.5, meaning that the distortion amounts on the average to a 50% increase of the reference field.

The existence of this high latitude, positive ΔB region and of the large negative decrease in the inner magnetosphere indicates that the inner magnetosphere is greatly inflated by a ring of particles, which is generally referred to as a ring current, even during quiet or only slightly disturbed periods.

The intensity of the quiet-time ring current is found to be greater by a factor of 2 or 3 than has generally been supposed and to vary considerably, even during slight disturbances. This suggests that the variation of the ring current intensity is directly related to activity in the plasma sheet region. Therefore, the investigation of the inflation of the inner magnetosphere is not only necessary for constructing a good theoretical model for the magnetosphere, which does not exist at the moment, but it is also an important part of the study of the dynamics of the magnetosphere.

DR. APEL:

Have you done any quantitative research on the ring current?

DR. SUGIURA:

We have not made any model calculations, but there is one aspect I did not mention because of the shortage of time; that is the high latitude, large ΔB region which results partially from this inflation of the inner magnetosphere and partially from the magnetopause surface current. This surface current may be much greater in the polar regions than we have thought.

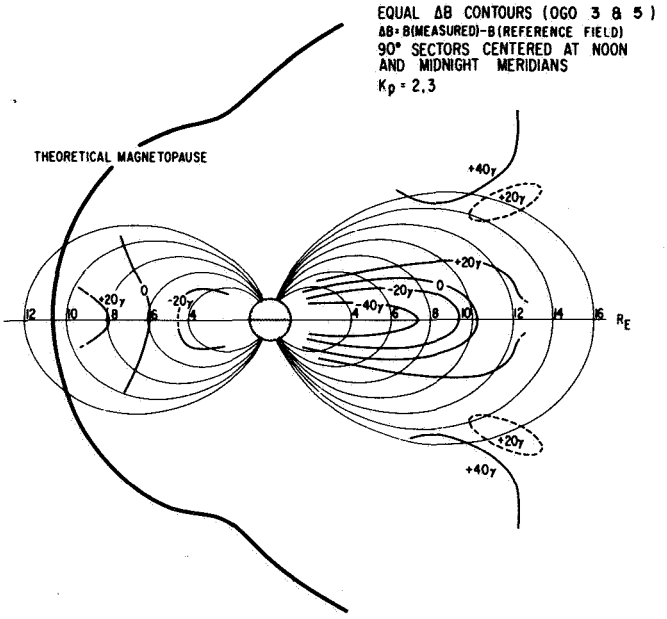


Figure 1

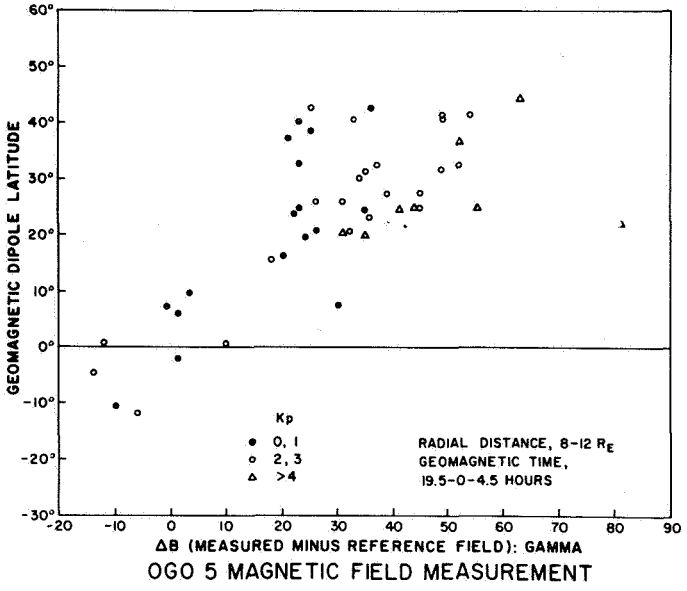


Figure 2

AURORAL PARTICLE INJECTION AND DRIFT

Robert A. Hoffman

CHAIRMAN:

Next we will hear from Dr. Robert Hoffman concerning auroral particle injection and drift.

DR. HOFFMAN:

In this talk I wish to show new information about the source location of electrons that cause the aurora in the auroral oval.

On the left side of Figure 1 is a view looking down at the north magnetic pole of the earth. To the right is midnight; the top, morning; to the left, noon. There are latitude circles of 60°, 70°, and 80° magnetic latitude.

When high-latitude magnetic observatories rotate through the midnight region they frequently observe sharp decreases in the horizontal component of the earth's magnetic field. These decreases, localized in the midnight region, are called magnetic substorms, or negative bays. At the onset of such a substorm the aurora overhead flares up.

The aurora is caused by the bombardment of the upper atmosphere by electrons with energies from about 1 to 10 kV. From measurements made by satellite-borne detectors in low-altitude polar orbits, it is observed that these electrons precipitate or bombard the atmosphere not only at midnight, but in a region extending from a few hours before midnight through the morning hours to local noon. What has not been known is where these electrons come from, and why. That is, are they locally accelerated in the region where they precipitate, or do they get there by some other means?

I will show that these electrons are injected onto magnetic field lines emanating from the earth in the restricted region of midnight during the time of a substorm and subsequently drift on closed field lines through the morning hours to local noon, where they are measured later precipitating into the atmosphere.

To prove this conclusively one must show that at midnight these electrons do exist at the time of a substorm, but they do not exist at midnight during magnetic quiet periods. On the contrary, at noon we must show that electrons do not exist at the time of a substorm, but do exist several hours

after a substorm, that is, the drift time for electrons to drift from midnight through morning to noon.

Figure 2 is another view looking down at the north magnetic pole of the earth. At 2252 U.T. on the 18th of September, 1967, there was a satellite pass through the midnight hours of the OGO 4 satellite, which contained an auroral particles experiment. About an hour later the satellite passed through the noontime hours.

The Kiruna, Sweden, magnetic observatory during this period of time was rotating through the midnight hours, and at about 1900 U.T. a substorm developed that lasted rather a long time. The midnight pass during the substorm shows large fluxes of electrons reaching up toward 10^7 electrons/ cm^2 sec sterad kV at the 7-kV energy measured, whereas at noontime the the satellite measured very low fluxes of electrons, about 1.5 orders of magnitude lower, hardly above the threshold level of the detector. This latter pass occurred when the midnight magnetic field was still depressed about 1000 gammas.

We see here during the time of the substorm at midnight large fluxes of electrons at midnight, but very few (essentially none) electrons at noon even 5 hours after the onset of the substorm.

In the interest of time I will not show the case of no particle precipitation at midnight during a magnetic quiet period.

Figure 3 is again a view looking down at the north magnetic pole of the earth. In this case, the College, Alaska, observatory was rotating through the midnight hours and observed a substorm commencing shortly before 14 hours U.T. We know from Figure 2 that electrons must have come into existence in the midnight hours at this time. About seven hours after the onset of this substorm a noontime pass of the OGO 4 satellite occurred, at which time the magnetic field in the midnight hours, now measured by the Kiruna, Sweden, observatory, showed no magnetic variation. However, at near noon the experiment on the satellite showed large fluxes of 7-kV electrons precipitating into the atmosphere, this being seven hours after the onset of the substorm.

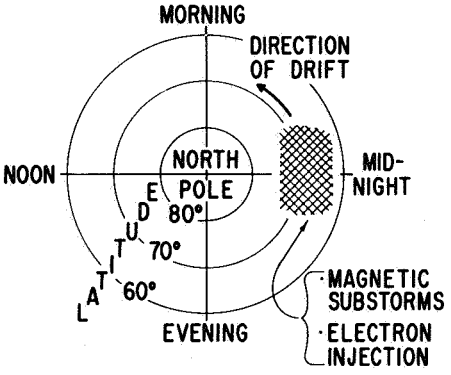
One concludes, then, that the source of electrons comes into existence in the midnight region at the time of the substorm. These electrons then drift in magnetic local time around the morning side, to noon, where they are measured some time later. Because the electrons are able to drift in longitude they must have been injected in the midnight region onto closed

field lines. And these closed field lines, for particles of the energies we are discussing, must exist over the morning hours to noon.

CHAIRMAN:

This represents I think the first time such an explanation of the behavior of auroral particles has been observed and made.

AURORAL ELECTRON INJECTION AND DRIFT



MIDNIGHT

Show electrons DO EXIST at time of a substorm.

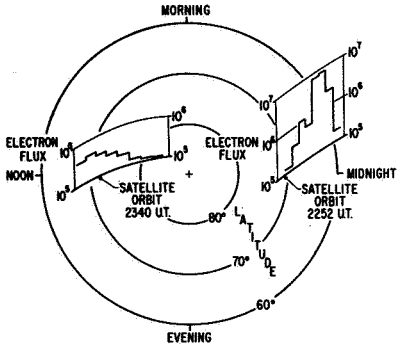
But DO NOT EXIST during quiet magnetic periods.

NOON

Show electrons DO NOT EXIST at time of substorm.

But DO EXIST several hours (drift time) after a substorm.

Figure 1



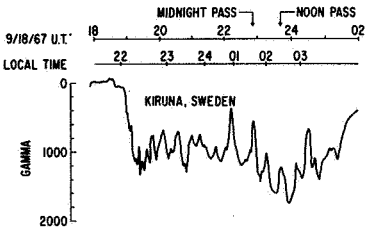


Figure 2

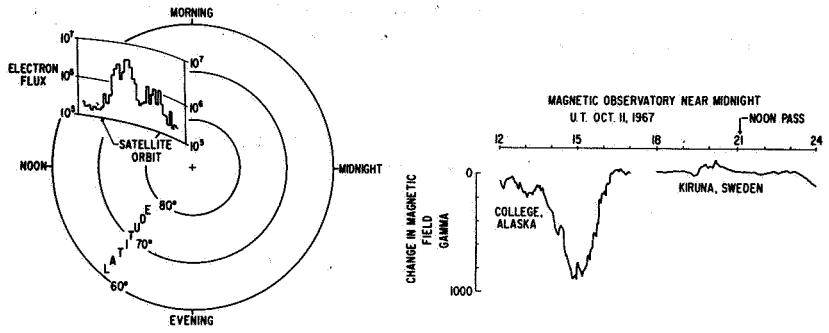

MORNING
ELECTRON FLUX
10^7
10^8
10^9
10^7
10^6
10^5
SATELLITE ORBIT
NOON
MIDNIGHT
80°
70°
60°
EVENING
MAGNETIC OBSERVATORY NEAR MIDNIGHT
U.T. OCT. 11, 1967
NOON PASS
12
15
18
21
24
0
1000
CHANGE IN MAGNETIC FIELD GAMMA
COLLEGE, ALASKA
KIRUNA, SWEDEN

Figure 3

OGO 5 MEASUREMENTS OF ELECTRONS NEAR THE MAGNETOPAUSE

Keith W. Oglivie

CHAIRMAN:

The next paper will be given by Dr. Keith Ogilvie concerning OGO 5 measurements of electrons near the magnetopause.

DR. OGILVIE:

In many theories of the magnetosphere large scale circulatory motions are proposed. Energy for driving these motions is supposed to come from the bulk motion of the solar wind. Plasma which interacts with the magnetosphere passes through the bow shock which divides the magnetosheath from the interplanetary medium, and then flows around the magnetosphere after being deviated through a large angle. Its streaming speed is reduced behind the shock, but it is still traveling at about 200 km/sec in the magnetosheath.

If we consider a fluid approximation there must be a boundary layer in the plasma between this flow and the very slowly moving plasma in the magnetosphere. The magnetopause boundary itself would then be one of the discontinuities which can exist in magnetohydrodynamics.

Parker and Lerch have predicted that fluctuations in the magnetic field are an inherent part of the structure of the magnetopause and that this boundary will be unstable, but since they assumed the field in the magnetosheath to be zero, their treatment is not a very realistic one.

The problem has also been examined theoretically by Eviatar and Wolf who conclude that a boundary 100 km thick would be sufficient to provide the viscous forces to drive, for example, the Axford-Hines magnetospheric model. The question of whether the magnetopause is a stable boundary layer in the hydromagnetic approximation, that is, either a rotational or a tangential discontinuity, is thus one of interest.

Observations of low energy electrons were carried out by a three-axis differential electrostatic analyzer spectrometer on the satellite OGO 5. This small device covered the energy range 10 eV to 10 keV in 16 differential steps. The three directional fluxes were sampled simultaneously, and it took 16 seconds to record a spectrum which was repeated every 23 seconds.

We are most interested here in the response of the 45-eV channel which accepts electrons with energies between 34 and 56 eV.

In Figure 1 we see how the flux behaves as the satellite traverses the magnetopause. The ordinate is the 45-eV flux in one of the electron spectrometer directions, and the abscissa is the universal time. The sharp increase is time-coincident to an accuracy of a few seconds with the magnetopause crossing as observed by the magnetometer of Paul Coleman on the same satellite, and with the change in the high energy particle flux observed by Harry West.

Figure 2 shows the observed energy spectra on both sides of the discontinuity. In the magnetosphere this spectrum is lower at low energies and higher at high energies than it is in the magnetosheath, as one would expect, since the magnetosphere contains trapped electrons, and the sheath, electrons heated by passage through the shock. The high energy flux values in these spectra are contributed by Harry West who measured electrons with energies between 80 keV and 1 MeV.

We have about 10 examples of this sort of crossing characteristic at geomagnetically fairly quiet times, with Kp around 3. For more disturbed times, of which we have a few examples, the data appear to be able to be represented as a combination of several of the more simple crossings.

These crossings all occurred on out-bound sections of orbits for which the satellite was on the dawn side of the magnetosphere in the forward hemisphere and at latitudes between 20 and 50 degrees.

The field magnitude changes, in general, across the discontinuity with both signs of change being observed.

There are five points in the rapidly rising section of the flux curve in Figure 1. If we assume this section to represent the boundary layer referred to, and further assume the discontinuity to be at rest, then the width is of the order of 100 km, or a few proton gyro radii. If the discontinuity is in motion, the thickness is correspondingly greater, but the simple nature of the transition in these cases may be interpreted as evidence that the boundary is at rest at this time.

We can draw a number of conclusions from our observations. First, the change in flux of low energy electrons appears to be an important indicator of a magnetopause crossing. Second, the magnetopause does not always seem to be moving. Third, if it is stationary, a thickness at the unstable region of about 100 km, in agreement with the theoretical ideas of Eviatar and Wolf, seems to be the correct order of magnitude.

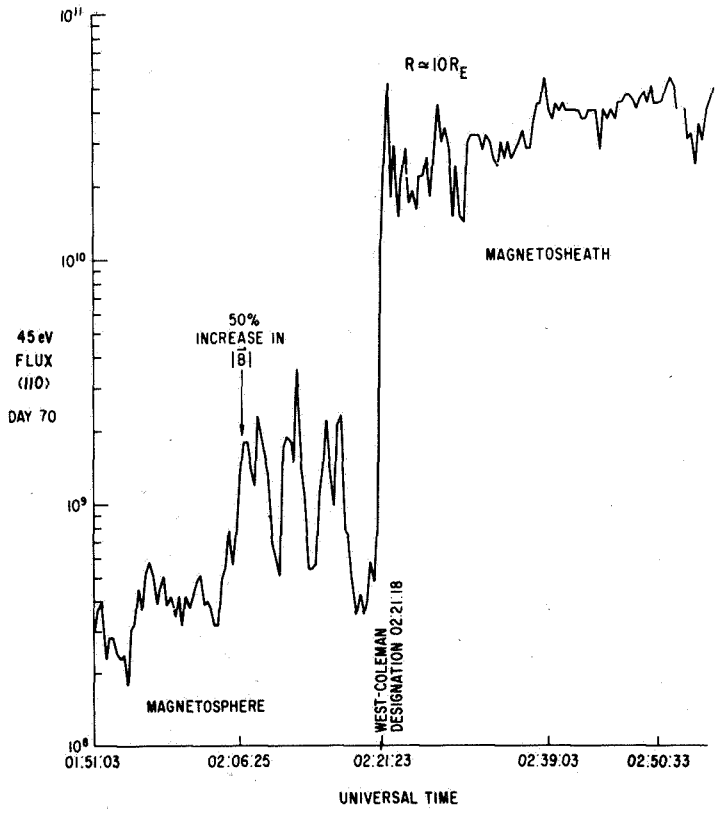


Figure 1

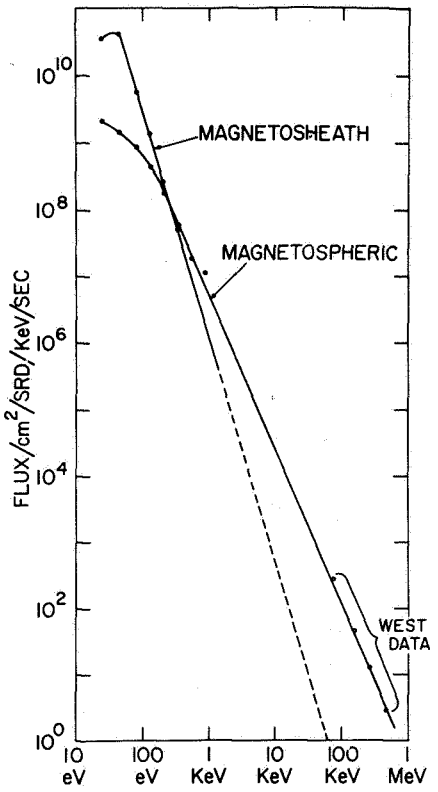


Figure 2 – OGO 5 energy distribution, day 70.

SHOCK METAMORPHISM OF BASALT AND THE LUNAR SAMPLES

Nicholas M. Short

CHAIRMAN:

The first paper on lunar studies will be presented by Dr. Nicholas Short.

DR. SHORT:

The most obvious features on the surfaces of the Moon and Mars are the circular depressions that cover much of the exposed areas. Most specialists in planetary processes identify such structures as impact craters, but a hard core of dissenters still contend that the majority of circular depressions are predominantly volcanic in character. The present consensus among planetologists favors the existence of both types but holds that impact dominates over volcanism. However, these conclusions are largely based on morphological comparisons, approaches that have proved ambiguous because two or more genetic processes can lead to similar outward appearances.

For several years, I have pointed out that the only satisfactory answer for the origin of lunar circular structures will come from rocks directly associated with these structures. If craters contain evidence of shock metamorphis, that is, unique, solid state melting effects caused by transient pressure waves in the hundreds of kilobar range, then an impact origin can be safely assumed.

Over 50 bona fide impact craters have been recognized on Earth, the older, eroded ones identified most readily by the presence of shock-metamorphic effects. Regrettably, no definite terrestrial impact structure, with the possible exception of one in India, occurs in basic igneous rocks, particularly basalt flows. If we assume that such basic rock types make up most of the Moon's outer layers, as indicated from Surveyor results and now confirmed by Apollo 11, then a study of some terrestrial impact crater in a similar rock type would be invaluable in preparing both the astronauts and principal investigating scientists for collecting and evaluating Moon samples. Despite a lack of impact craters in terrestrial basalt, there is an alternative.

In 1962, a small nuclear explosion, the Danny Boy event, produced a 70-m-wide crater in basalt at the Nevada test site (Figure 1). Such nuclear

craters duplicate many aspects of impact craters and contain rock that has proved to be quite pertinent to current studies of the lunar rocks. The Danny Boy basalt consists of a calcium-rich feldspar, a magnesium-iron silicate called olivine, and the iron oxide, magnetite. At pressures up to about 300 kb, only microfracturing of the feldspar and olivine occurs (Figure 2a). At higher pressures, up to about 450 kb, the feldspar is gradually converted to an isotropic phase that appears dark in polarized light (Figure 2b). Over this 450-kb interval, shock damage in basalt is far less obvious and spectacular than in granitic rocks. This means that, on first inspection, the bulk of basaltic lava samples may not show direct indication of having been shocked. Above 450 kb, the changes in Danny Boy basalt are abrupt and dramatic.

Figure 2c shows a stage in which the feldspar had melted but did not mix effectively with other constituents. At even higher pressures, olivine granulates and begins to disperse into the melt, reacting with it to form one or more unusual minerals (Figure 2d). However, because of the extreme rapidity of the transformation process (just fractions of a second), the original textures of the rock are preserved except for the formation of gas-expansion cavities. When these observations are applied to the lunar samples, the results are explicit.

Signs of shock are abundant in the lunar soil and fragmental rocks and are possibly present in the crystalline rocks. The object in Figure 3a is a feldspar crystal fragment containing possible planar features. These are mechanically induced discontinuities whose formation is known to be due exclusively to shock compressions greater than approximately 100 kb.

Figure 3b shows what were once fragments of feldspar sand, now shock-lithified into what I call "instant rock." Figure 3c shows shock-induced deformation bands in pyroxene, and Figure 3d shows a sphere of glass that is almost certainly a result of impact melting of some lunar rock material.

These are obvious examples of shock damage that are analagous to the Danny Boy basalt, but most fragments in the lunar microbreccias appear unshocked. This is precisely the prediction of the Danny Boy study, which indicated a strong resistance to shock pressures up to about 300 kb. Thus, most basaltic rocks involved in impact cratering will not show obvious signs of damage. Overall shock damage in Danny Boy basalt differs from that found in shocked Moon rocks mainly in the absence of long-term alteration effects (e.g., recrystallization) and in its nonfragmental character.

One of the major scientific tasks of the remaining Apollo missions will be on-the-spot examination of the voluminous shock ejecta deposits (perhaps 2 km thick). This should be accompanied by careful selection of samples for detailed analysis on Earth by experts in the field of shock metamorphism.

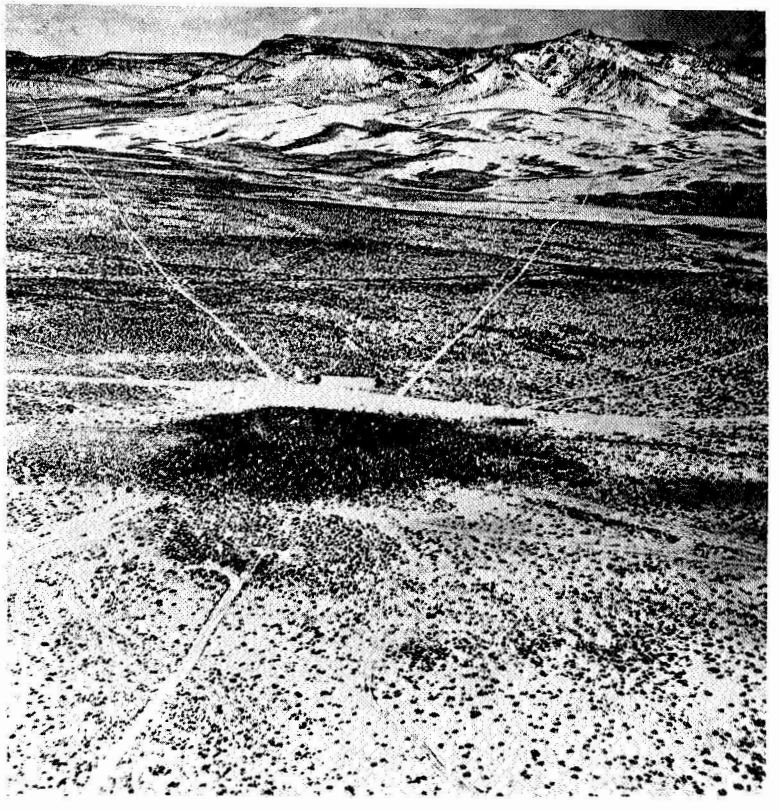


Figure 1

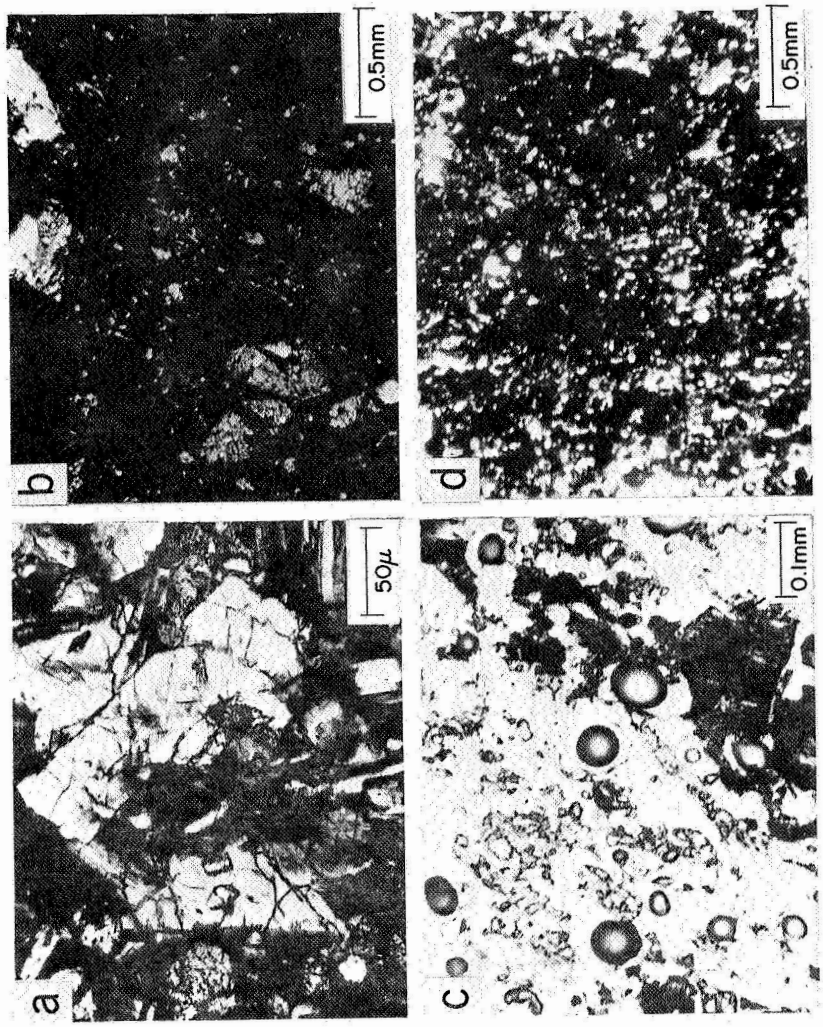


Figure 2

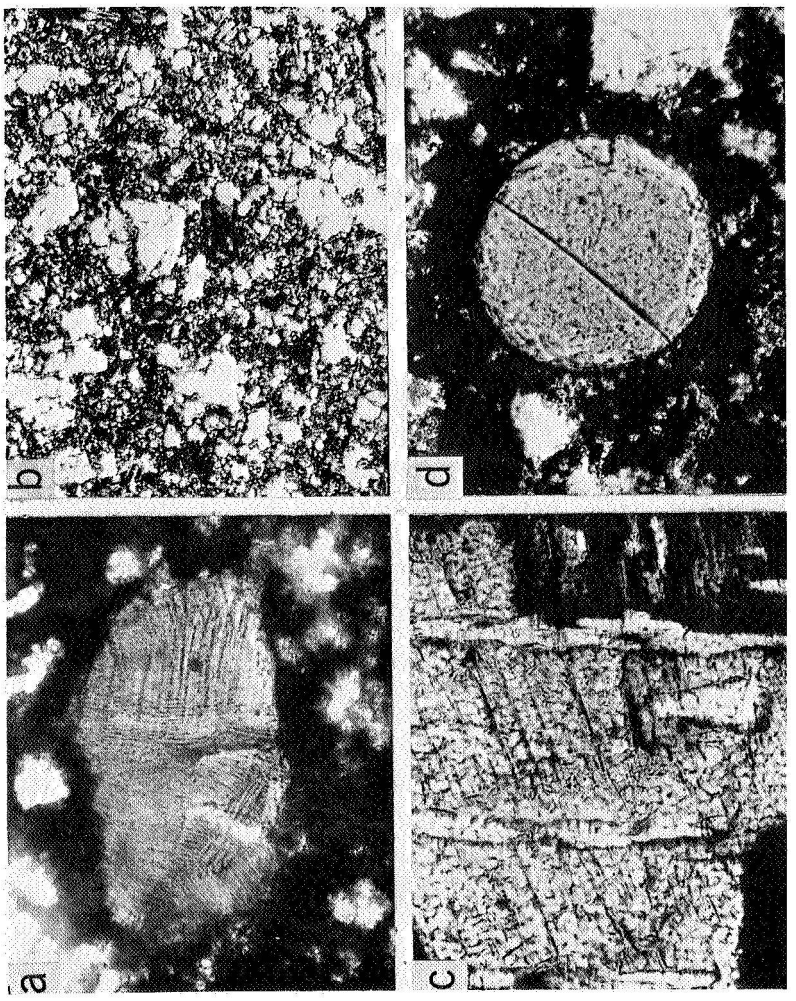


Figure 3

LUNAR MASCONS DETERMINED FROM THREE APOLLO MISSIONS

James P. Murphy

CHAIRMAN:

We will move on next then to Mr. James Murphy, who will tell us about lunar mascons from three Apollo missions.

MR. MURPHY:

The effects of mascons, or mass concentrations, have been noticed in the tracking data on all manned lunar missions.

The first of these was Apollo 8. The tracking data from this mission have been used to improve our knowledge of the lunar gravity field. In Figure 1 we see the orbit determination range rate residuals as the spacecraft passed over the Sinus Aestuum region on the lunar surface in its retrograde orbit. We have these range rate residuals for the first orbit in the upper curve, and for the seventh orbit in the lower curve.

The first orbit was more elliptical than the seventh orbit, and the spacecraft passed over this region at 310 km. In the seventh orbit, the spacecraft passed over this region at 112 km. Notice that the rise in the range rate residuals is steeper for the closer orbit than it is for the more distant one.

Since the range rate was positive at this time, and since the range vector and the radius vector were very nearly colinear, the rise implies that the spacecraft was accelerated toward the lunar surface at a greater rate than was expected.

Now from Newton's laws we know that the acceleration on an object of negligible mass due to an external mass is equal to a constant times the mass of that object divided by the square of the distance between the two.

In Table 1 we have the values of the observed acceleration at the two heights obtained from taking the slope of the two curves for the higher and the lower orbits; we also have the heights shown there. So we may apply Newton's laws twice and obtain two equations in two unknowns, which are easily solved.

Table 1

Apollo 8 Passage
over Sinus Aestuum

Orbit Data	Elliptical Orbit	Circular Orbit
Longitude	−8°	−8°
Latitude	10°	10°
Acceleration	0.15 mm/sec^2	0.56 mm/sec^2
Height	310 km	112 km

The mass, then, for the mass concentration turns out to be 3.8×10^{20} g, or 5 millionths of the mass of the moon, or five micromoons, and the depth turns out to be 100 km. The mean radius of the moon is 1738 km.

This represents the first solution for both the mass and the depth of the mascon. Previous investigators had accelerations at only one height, or nearly the same height, and were able to obtain masses only after assuming depths.

In Figure 3 we see range rate residuals for the tenth orbit of Apollo 10. In this case the spacecraft passed directly over the mascons at Mare Smythii and Sinus Medii and near another one at Mare Orientale. As the spacecraft exits occultation, the mascon at Mare Smythii is between the observer on earth and the spacecraft. After it passes the mascon, the spacecraft is between the observer and the mascon. Hence we see a rapid drop and rise in the range rate residuals as the mascon quickens and then retards the movement of the spacecraft along its orbit. The texture of the curve at the center is similar to the Sinus Aestuum region for Apollo 8. In the case of the passage near Mare Orientale, the events take place in opposite order from that for Mare Smythii, and hence we have a rise and then a fall in the range rate residuals.

From the Apollo 10 analysis, three mascons were obtained. When these mascons were added to the operational lunar field, together with the mascon obtained from the Apollo 8 mission, a noticeable improvement was obtained in the range rate residuals for Apollo 11. The standard deviation in the residuals for the range rate tracking data was improved for a typical orbit of Apollo 11 from 2.2 to 1.2 cm/sec.

The tracking data from the Apollo 12 mission have been used to obtain the mass and depth of mascons in Mare Nectaris and Oceanus Procellarum.

When these two mascons, together with the mascon at Mare Smythii, were added to the Apollo 12 operational gravity model, considerable improvement to the fit in the data was again achieved. For the third lunar orbit the standard deviation in the range rate residuals was reduced from 3.31 to 1.11 cm/sec. For the range data this reduction was from -166 to -67 m. This value of -67 m is of a comparable size to the known error in the lunar ephemeris used for this mission at this epoch.

MEMBER OF THE AUDIENCE:

Mascons have not been determined to be underneath any feature other than a ringed mare, have they?

MR. MURPHY:

Well, I do not think that is so because mascons have been obtained at Sinus Medii and Sinus Aestuum, and they are not maria at all. The speculation was that at one time they might have been maria and they were distorted by further impact. Furthermore, the mascon that we obtained in the Ocean of Storms does not correspond to a ringed mare either.

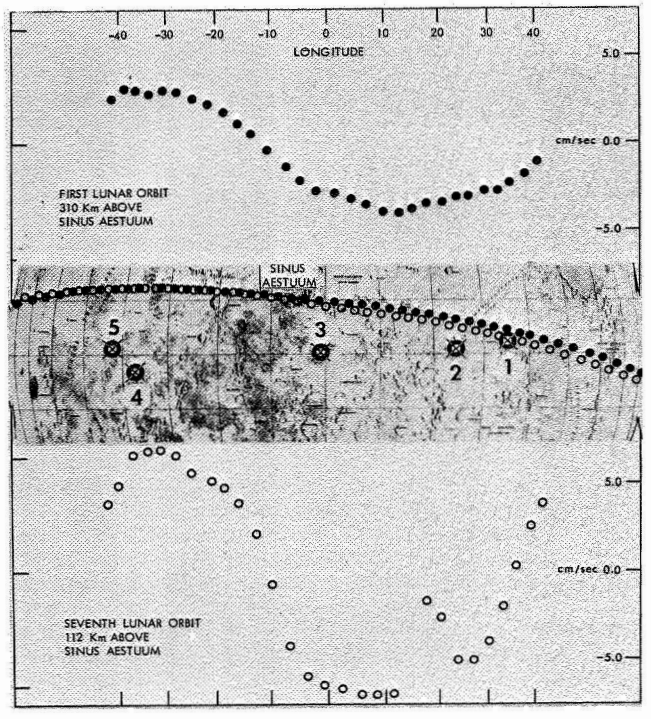


Figure 1 – Apollo 8 range rate residuals.

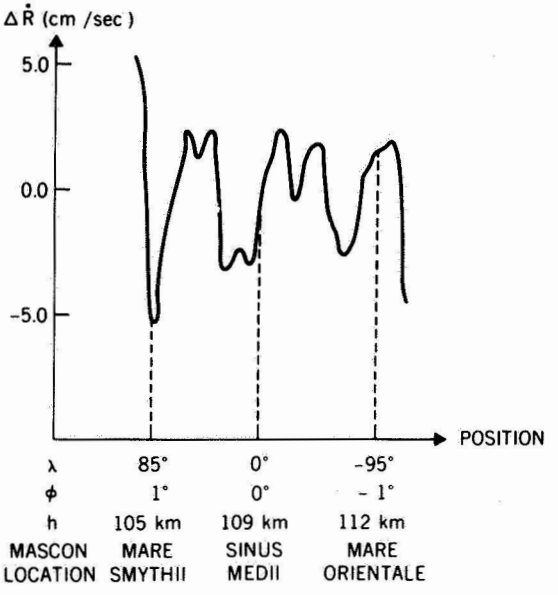


Figure 2 – Apollo 10 range rate residuals.

A DIFFERENT LOOK AT MASCONS

Donald S. Woolston

CHAIRMAN:

I am going to interrupt the regular program now to introduce a post-deadline paper, you might say. This is one which came to my attention after the program was essentially put together. I think it is a matter of sufficient interest, as I think you will see, that it deserves to be introduced.

It is being given by Mr. Donald Woolston, and the title is "A Different Look at Mascons."

MR. WOOLSTON:

By way of contrast to what Mr. Murphy has just told you about mascons based on life in the real world, we have taken another look through some studies based on generated tracking data. Here one creates his own perfect world and is then able to perform controlled numerical experiments.

We believe that the mascon problem was detected because, for vehicles orbiting the moon, there arose for the first time the opportunity for nearly continuous tracking coverage. This gave a very detailed time history of the vehicle's response to the perturbing force field and would clearly show up any anomalies.

It has been our suspicion for some time that, when orbiting tracking stations provide continuous coverage of earth satellites, something akin to earth mascon problems may arise. The present study of the response of various orbit determination schemes to errors intentionally introduced in the earth's force field was begun with the hope of gaining some understanding of these problems before we were actually confronted with them.

Figure 1 indicates the operation of a typical orbit determination scheme. We have selected a particular reference orbit, designated Orbit I (Ref), and have used it to generate a fictitious set of measured observations as seen from some tracking station. These measurements will be my data and are to be thought of as actual tracking data. Please forget that in reality I computed them.

To show how my orbit determination program performs, I want now to take the very same reference orbit and introduce an error in the force field.

In Figure 1 this is designated Ref. Orbit w/error. With this error present, we have produced a set of computed observations. If we now take the difference between the original measured observations and these computed observations, we arrive at a set of what I want to call preliminary scan residuals.

To this point, no orbit correction at all has been attempted; these preliminary scan residuals simply show the effect of the presence of the force field error on our reference orbit.

These residuals are fed into the orbit determination program which processes them statistically, typically by least squares techniques, to arrive at an orbit which best fits the tracking data. Normally this is an iterative process and requires several passes of the data through the orbit determination program. For simplicity I have shown a single pass, and I am calling my best fit Orbit II.

Next, I want to take Orbit II and, with the very same force field error present, compute a new set of observations. If now I take these new computed observations and subtract them from the original measured ones, I arrive at what I want to call a set of final scan residuals. These will show the effect of the presence of the force field error on my best fit orbit. In the examples which follow, I will compare the preliminary scan residuals with the final scan residuals.

Figure 2 illustrates the case of an error introduced in the mass of the earth. My data, the fictitious measured observations, were generated with the currently accepted value of the earth mass parameter μ, labeled observed μ. To generate the computed observations, an error was introduced in the fourth place of μ; this is the value labeled computed μ.

The two curves shown here give range rate residuals in kilometers per second as a function of time. I apologize for the scale, which is the output of the computer program, but we will be concerned only with the magnitudes of the numbers involved.

The curve in Figure 2(a) gives the preliminary scan residual pattern and shows simply the effect of the earth mass error on the reference orbit. Note that the largest residuals occur near perigee. The curve in Figure 2(b) gives the final scan residual pattern and shows the effect of the error in earth mass on the best fit orbit.

Now if the orbit determination program had exactly recaptured the true, or reference, orbit, both curves would be identical. Note, however, that while the magnitude of the residuals has been greatly reduced, a spurious frequency has been in the residual pattern.

A final scan residual pattern such as this is used by orbit analysts to deduce the physical source of errors in the force field. We suspect that a satellite geodesist looking at the curve in Figure 2(b) would be more likely to associate this particular pattern with an error in some coefficient in the oblateness potential rather than in the earth mass, as is actually the case. This implies to us, then, that there is some need for caution in the analysis of the frequency content of residual patterns.

For the final example, we restored the earth mass parameter to its proper value and introduced an error in the shape of the earth. The data—that is, the measured observations—were generated from a reference orbit with the currently accepted model of the earth oblateness potential. To obtain the computed observations, however, we zeroed out the oblateness potential and assumed a spherical earth. Again we found that the final scan residual pattern contained spurious frequencies introduced by the orbit determination process.

Figure 3 shows a comparison of the results for this earth shape error case with lunar orbiter residuals. Figure 3(a) presents our final scan residual pattern and shows the effect of the earth shape error on the best fit orbit. Figure 3(b) presents the lunar orbiter residuals. We submit that the results are strikingly similar.

The curve in Figure 3(b) and others like it were analyzed by Muller of JPL to arrive at the mascon concept. We are not at this point in a position to affirm or deny the existence of mascons. Certainly there is strong evidence in support of them—Jim Murphy's results, as well as JPL's work showing correlation with lunar surface features. All we are saying is that our results do indicate some need for caution.

Our study, which is still quite preliminary, simply indicates that several error sources which are unrelated to mascons can produce residual patterns much like those from which mascons were deduced. We do suggest that further investigation is warranted.

MEMBER OF THE AUDIENCE:

Is there any explanation of why they should have located the things right over the maria?

MR. WOOLSTON:

Well, in the case of Lunar Orbiter, which had an eccentricity of about 0.5, I think that these residual patterns were commonly associated with perilune, the closest approach to the moon, which I think happened to coincide with the maria.

MEMBER OF THE AUDIENCE:

I would like to make the point that, for the earth oblateness potential, the zonal harmonics are usually obtained by the analysis of many arcs, that is, by studying the long term variation in the angular variables.

MR. WOOLSTON:

That is true. We have looked just at a single pass. We certainly plan to take a further look. Our people were at JPL on Monday, and I understand that on the basis of our results, they plan to take another look also. But as I say, it is a rather preliminary result at this point.

DR. HALL:

I do not think it is related to your point, but I was tempted to mention a moment ago on the earlier paper: there is also the ambiguity on the interpretation of any external geopotential observation, because there is no unique solution. There is a very large variety of possibilities which can account for any potential.

CHAIRMAN:

What we need is an infinite series, a complete orthonormal set of mascons.

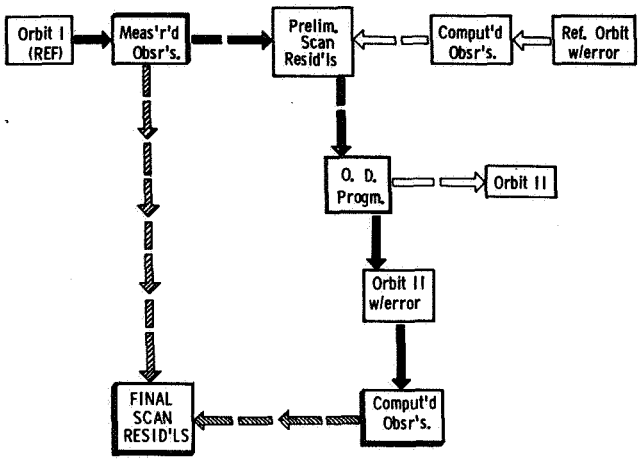


Figure 1

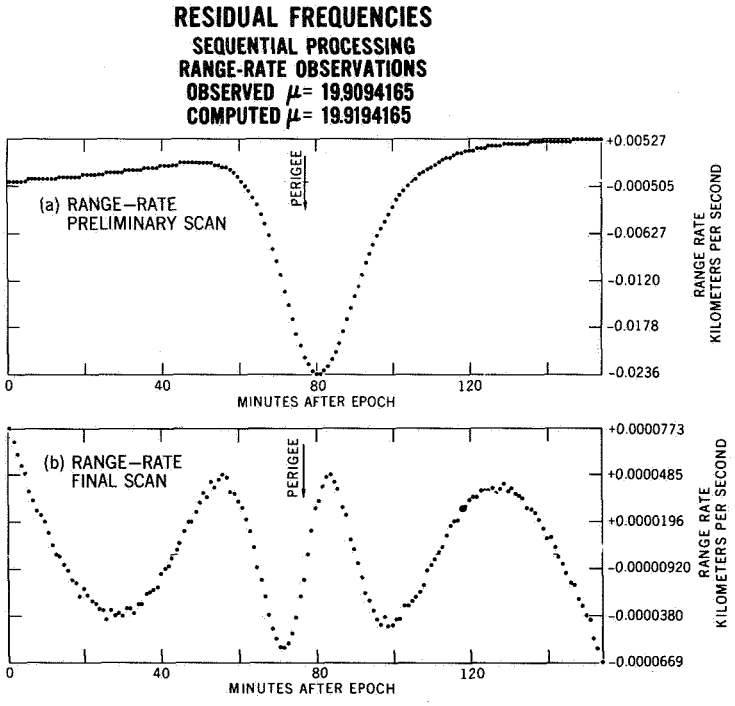


Figure 2

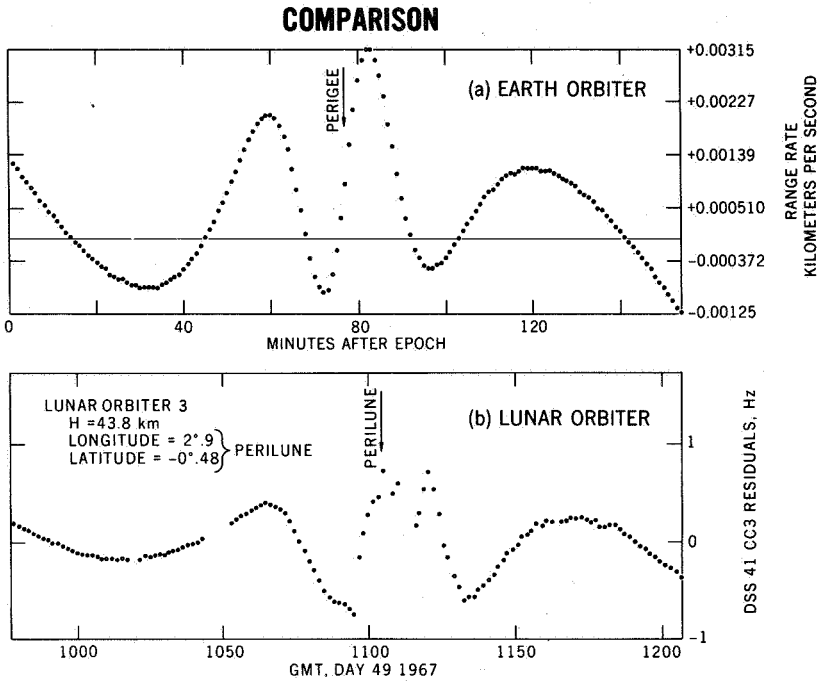

COMPARISON
(a) EARTH ORBITER
PERIGEE
+0.00315
+0.00227
+0.00139
+0.000510
-0.000372
-0.00125
RANGE RATE
KILOMETERS PER SECOND
0
40
80
120
MINUTES AFTER EPOCH
LUNAR ORBITER 3
H =43.8 km
LONGITUDE = 2°.9
LATITUDE = -0°.48
PERILUNE
PERILUNE
(b) LUNAR ORBITER
1
0
-1
DSS 41 CC3 RESIDUALS, Hz
1000
1050
1100
1150
1200
GMT, DAY 49 1967

Figure 3

THE STRUCTURE OF THE MARTIAN ATMOSPHERE FROM MARINERS 6 AND 7

S. Ichtiaque Rasool

CHAIRMAN:

Next, we will hear a discussion of the structure of the Martian atmosphere by Dr. S. Ichtiaque Rasool.

DR. RASOOL:

The highly successful Mariner 6 and 7 missions to Mars have provided detailed information, for the first time, on the structure of the Martian atmosphere. The encounter with Mars took place about three months ago, and I wish to report some of the first results obtained concerning the structure of the lower atmosphere of Mars, namely the temperature and pressure determined by the S-band occultation experiment that was conceived by Dr. Kliore at the Jet Propulsion Laboratory and in which I participated as a coexperimenter.

The principle of the experiment is actually very simple, and the beauty of the experiment is that it does not require any extra instrumentation on the spacecraft; we made use of the highly sophisticated telemetry system that was already on board.

The only requirement was that the spacecraft go behind the planet, as seen from the Earth. Figure 1 shows the trajectories of Mariners 6 and 7. Just before the spacecraft disappeared and just after it reappeared, the signals passed through the atmosphere of Mars above a given point on the planet. Because of the presence of atmosphere, the signals are modified in both phase and frequency, and this can be recorded very accurately on Earth. The changes in the phase are proportional to the refractivity of the atmosphere, and, by assuming the composition, we derived the density as a function of height above the occultation points. Knowing the density distribution as a function of height, we can derive temperature and pressure quite accurately.

By using this technique, we now have distributions of temperature and pressure at four locations on the planet: one at the equator and another near the pole on the day side and one in the midlatitudes and another near the pole on the night side. Figure 2 shows the results. These are the four temperature profiles as functions of pressure, computed for a composition of 100% CO_2. The measurements were made at altitude intervals of 2 km. The

last point in each profile is the cut-off point, so it gives the pressure at the surface.

Apart from the fact that these profiles will eventually be used to derive the dynamics of the Martian atmosphere, there are a few interesting points that have emerged already. The first of these concerns the surface pressure. At three widely separated points on the planet, the pressure is between 6 and 7 mb. At the fourth point, however, the pressure is only 3.5 mb. If this is interpreted as an elevation difference of the occultation point, a calculation gives a height of about 6 or 7 km at this point (58° south latitude).

We knew from earlier radar measurements that the equator of Mars has a rugged topography, but no information was available about the heights of these features, especially near the poles. This region, which we find now to be high, was known as Hellespontus Depressio.

The second important piece of information that we obtain from Figure 2 is the temperature range in the atmosphere. We knew from earlier ground-based measurements that the temperature at the surface of Mars varies a great deal between day and night and between the equator and the poles (100° K differences between the noontime temperature at the equator and the nighttime temperature at the poles), and that is what we found. But all theoretical calculations indicated that because the atmosphere of Mars is very thin—the pressure is only 6 to 7 mb—the temperature in the atmosphere should not vary very much. The maximum estimated variation between the daytime-equatorial and nighttime-polar temperatures is 30° K. Yet, we found a difference of about 100° K in the middle atmosphere. The accuracy in the lower atmosphere is about ± 15° K, and this range decreases at higher altitudes where there is less atmosphere.

It is important to note that, at the equator, the temperature appears to be uniform with height, which is completely different from what we estimated theoretically. This probably implies that the opacity of the atmosphere is much higher than we had thought. Either the opacity of the CO_2 is higher than what we had thought it to be, or there is dust present in the atmosphere that absorbs the infrared and solar radiation.

The third point of interest is the temperature near the north polar cap at night. The temperature is very close to the condensation temperature of CO_2 at this point. It is even close to the condensation point near the ground. This is an independent argument that the polar caps are made of dry ice.

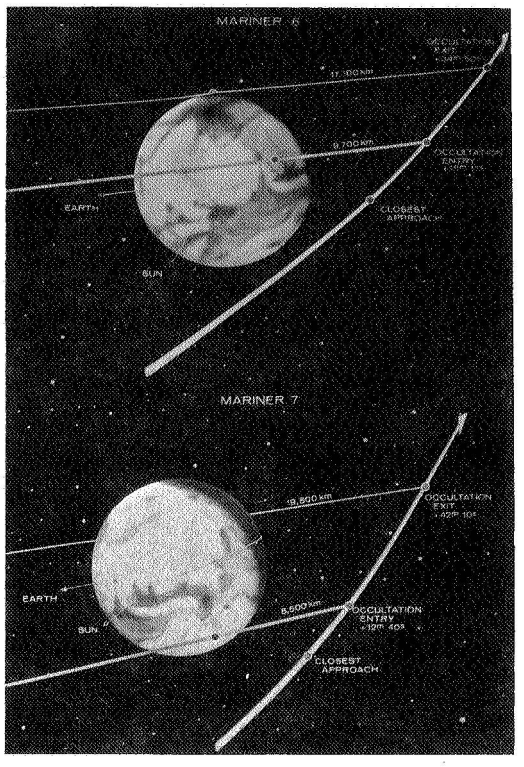


Figure 1

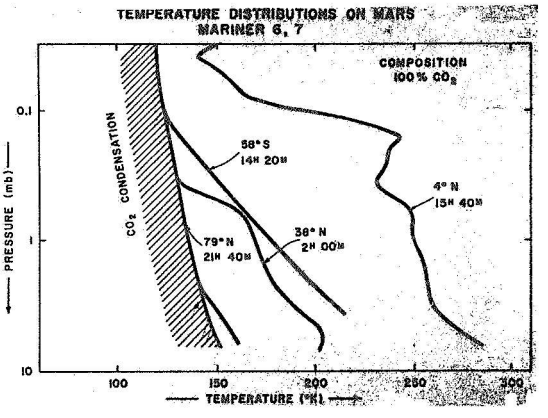


Figure 2

THE CONTRIBUTION OF ATMOSPHERIC AEROSOLS TO THE MARTIAN OPPOSITION EFFECT

Jaylee M. Mead

CHAIRMAN:

The next paper concerns the contribution of atmospheric aerosols to the Martian opposition effect and will be given by Mrs. Jaylee Mead.

MRS. MEAD:

Observations of Mars by Brian O'Leary during the 1967 opposition show on opposition effect, that is, a nonlinear surge in brightness as the planet approaches a phase angle of zero degrees.

The opposition effect is shown in Figure 1 for five colors–ultraviolet (U), blue (B), visual (V), red (R), and infrared (I). Reflectivity, or albedo, is plotted on a logarithmic scale on the right-hand side and an equivalent magnitude scale on the left. It covers a range of 2.5 magnitudes, a factor of 10. Across the bottom is phase angle α, which is the angle at the planet between the sun and the observer.

The Martian observations are shown as a solid line; they were made between 1.2° and 8° of phase angle. These observations have been fitted to the linear phase functions for larger phase angles given by de Vaucouleurs.

As O'Leary and Rea pointed out, the opposition effect is much more pronounced in the blue than it is in the red. In other words, there is a much greater departure from the linear extrapolation in the ultraviolet and blue than there is in the red and infrared. In addition to this greater opposition effect in the ultraviolet, we see that the reflectivity, or albedo, is much higher in the infrared.

The question is, what could cause this increased opposition effect in the blue and ultraviolet? It could be primarily a surface effect in that the surface could have a much greater reflectivity at shorter wavelengths, or it could be due to the Martian atmosphere.

The purpose of this study was to investigate the contribution which atmospheric aerosols might make to the Martian opposition effect. To do this I have made light-scattering calculations using the Mie scattering theory, which is single scattering by spherical particles. I have done this

for a number of substances, such as ice, water, and solid carbon dioxide, which have been identified as constituents of the Martian atmosphere.

I have also made calculations for highly absorbing substances such as limonite, a ferric oxide which has been suggested as a possible Martian surface covering. Various submicron particle-size distributions have been applied, and the integrated intensities for these aerosols have been computed as a function of phase angle for the wavelengths mentioned.

Table 1 summarizes the results of the light-scattering calculations for a variety of indices of refraction and particle-size distributions of aerosols. Neither the highly absorbing substances, such as limonite, nor substances having an index of refraction less than 1.5–which include water, ice, and solid carbon dioxide–show an opposition effect. However, calculations for refractive indices from 1.55 to 2.0 do show a definite increase in brightness at small phase angles with little or no enhancement at larger phase angles.

Table 1

Results of Light-Scattering Calculations

Index of Refraction n	Opposition Effect
$1.20 \leq n \leq 1.50$	No
$n = 2.23–0.669\ i$ and other indices with large absorption	No
$1.55 \leq n \leq 2.0$	Yes

Having found that aerosols with a refractive index of 1.55 or greater could produce an opposition effect, we next introduced a surface brightness function into the model. The results are shown in Figure 2. The broken line indicates the Martian observations. At longer wavelengths where the Martian albedo is higher and where surface markings are more clearly visible, it is reasonable to assume that the observed brightness is due primarily to the surface and that brightness contributions by atmospheric aerosols would be negligible.

Therefore, we fit a surface brightness function to the observations in the infrared and assumed that it would have the same shape at all other wavelengths; only the albedo would change. The thin, lower curve in Figure 2 represents the assumed surface function for each of the wavelengths shown.

We then added to this the effect of the aerosols. The calculated brightness of the surface plus submicron aerosols having a refractive index of 1.75 is shown as the upper, heavy solid curve for each wavelength. At shorter wavelengths, where the albedo and surface contrast are greatly reduced, we see that atmospheric aerosols can play a significant role. The calculated phase curves are seen to be in reasonable agreement with the observations. This demonstrates that a small contribution by atmospheric aerosols having the desired optical properties can provide the observed opposition effect at small phase angles in the ultraviolet, but, at the same time, they can have a very negligible effect in the infrared.

MEMBER OF THE AUDIENCE:

What kind of aerosols did that?

MRS. MEAD:

Well, that is a good question. Actually, some good candidates are shown in Table 2, which is a list prepared by Coffeen. The most promising are semi-transparent minerals. Unfortunately, we know only the real part of the indices of refraction because there are often impurities in the natural substances which make it harder to determine a definite or consistent absorption.

I might add that we would like more measurements of the refractive index of solid carbon dioxide because the ultraviolet measurements which I have used have been made by only one group. Also, it would be desirable to have these measurements made at lower temperatures.

CHAIRMAN:

That had all the earmarks of a well-organized plant. Well done, Jaylee.

Table 2

Refractive Indices

Liquid		Solid	
CO_2	1.20	H_2O	1.31
NH_3	1.33	NaF	1.34
H_2O	1.33	N_2SO_4	1.43
D_2O	1.33	CaF_2	1.43
CH_3CH_2OH	1.36	KCl	1.49
SO_2	1.41	NaCl	1.54
H_2O_2	1.41	$C_{12}H_{22}O_{11} \cdot H_2O$	1.54
C_3O_2	1.45	SiO_2 (crystalline quartz)	1.55
CCl_4	1.46	$Mg(OH)_2$	1.57
C_6H_6	1.50	$CaCO_3$ (calcite)	1.60
$C_6H_5CH{:}CH_2$	1.54	$CaSiO_3$ (wollastonite)	1.63
CS_2	1.63	NH_4Cl	1.64
		MgO	1.74
		$AgNO_3$	1.76
		CaO	1.84
		C (diamond)	2.42

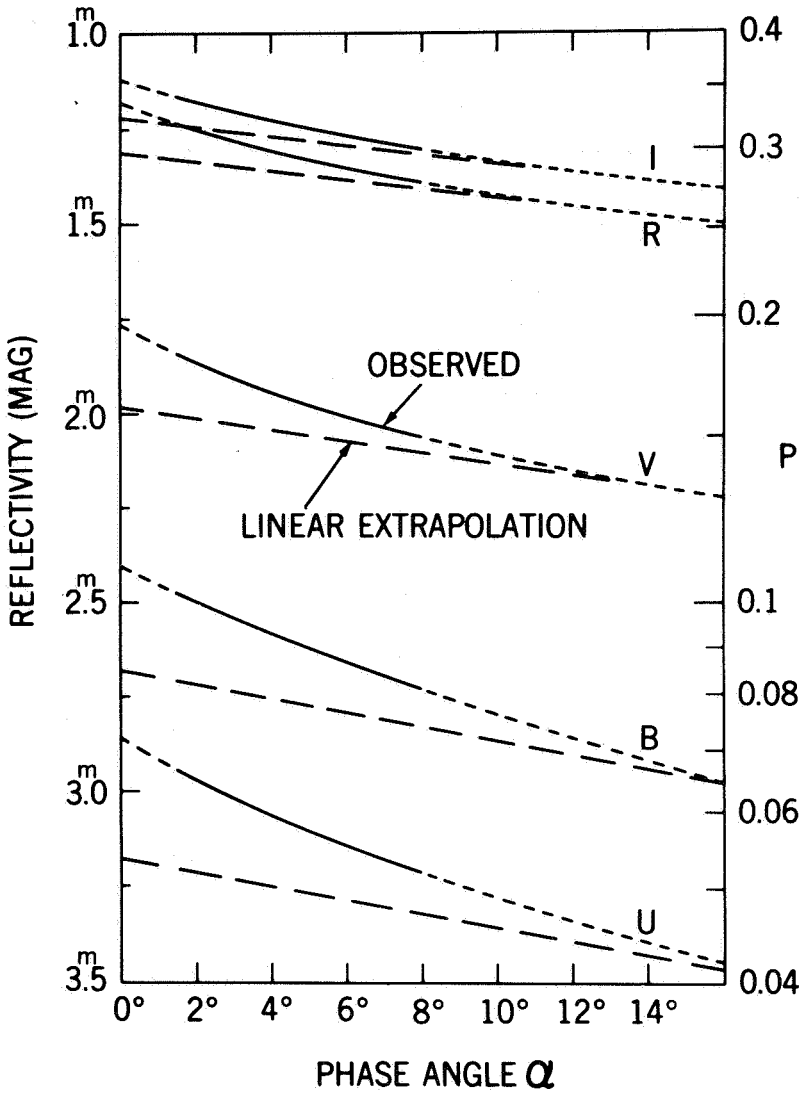


Figure 1–Martian opposition effect (after O'Leary and Rea, *Icarus* 9:405-408, 1968).

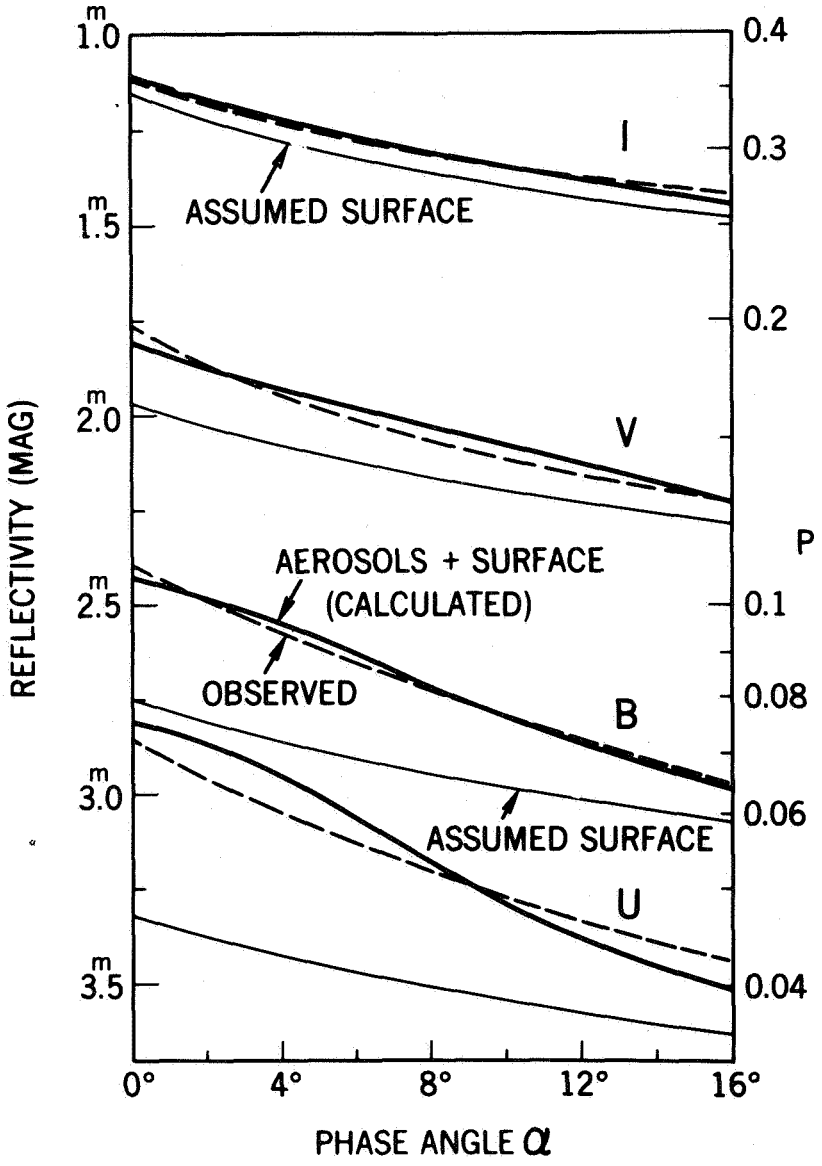


Figure 2 – Comparison of model with Mars observations for submicron aerosol size distribution with $n = 1.75$.

PHYSICAL PROPERTIES OF THE VENUS IONOSPHERE

Jay R. Herman

CHAIRMAN:

The final paper in this morning's session will be given by Dr. Jay Herman, on physical properties of the Venus ionosphere.

DR. HERMAN:

After all these talks I would not want us to discriminate between the planets, so here we go.

During the 1967 fly-by of Venus, Mariner 5 obtained a single cross section of the daytime ionosphere at a point about 45° from the subsolar wind point. These data are shown in Figure 1.

The heavy curve on the bottom of Figure 1 represents the data obtained as published in *Science* for the 1967 Mariner fly-by, showing the sharp cut-off in the topside and the electron density peak. The rest of the data shown refer to the night side, which I am not going to discuss in this talk.

The surprising feature contained in the electron density profile was an abrupt termination of the electron density, which is called the ionopause, at an altitude of about 500 km; that is, in a distance of about 50 km the electron density goes from 10^4 to about 10 per cm^3.

This electron density profile and the few parameters that can be inferred directly from it are the basis of our experimental knowledge of the Venus ionosphere.

To learn more, we have constructed a theoretical model of the ionosphere by simultaneously solving the equations of heat transport for the electron, ion, and neutral gases, along with momentum and chemical equations for the charged particle densities.

Of all the ionospheric data, the abrupt termination seen in the electron density profile is the essential starting point for the construction of a theoretical model. When this model is brought into approximate conformity with as much of the data as is possible, we are able to obtain the first meaningful estimates of the electron and ion temperatures and the horizontal magnetic field strength in the topside ionosphere.

The key to performing the analysis lies in the interpretation of the ionopause.

In addition to the ionospheric parameters Mariner 5 also measured the important characteristics of the solar wind as it passed by Venus. Shown in Figure 2 are the electron density (3 per cm^3) and the associated solar wind velocity (590 km/sec) measured by the spacecraft just prior to the occultation experiment. The point where the previous profile was obtained is shown in Figure 2 at 45° below the subsolar point.

The supersonic flow of solar wind interacts strongly with the planet to form a bow shock observed at a distance of about 50,000 km, as represented by the solid line on the outside. Closer in, the solar wind and its magnetic field start to pile up on the top side of the ionosphere, because the magnetic field frozen into the solar wind cannot penetrate into the highly conducting ionosphere. Thus, the solar wind flow must turn and sweep around the planet, while at the same time carrying away the topside ionosphere. This process can continue until the internal pressure of the ionosphere is just able to balance the solar wind pressure of 8.78×10^{-9} dynes cm^{-2}.

This pressure balance, combined with the measured electron density of 10^4 per cm^3, just below the ionopause, enables the boundary conditions for the temperature to be determined. This would be sufficient to specify the thermal structure throughout the ionosphere, if it were not for the presence of the horizontal magnetic field within the ionosphere.

The horizontal field arises from two sources: First, as the solar wind and its frozen-in magnetic field pile up on the ionopause, some of the field is able to leak inside, since the ionosphere is not a perfect conductor. And, second, Venus can be viewed as a conductor embedded in a moving magnetic field, which means that an electric potential appears across the planet. This potential forces currents to flow through the ionosphere parallel to the surface of the planet, thus generating a horizontal magnetic field.

Once a horizontal magnetic field is present, it means that the electron and ion thermal conductivities are reduced, and the temperature and topside plasma pressure must increase.

From the requirements imposed by the pressure balance we have been able to determine that the strength of the magnetic field in the topside Venus ionosphere is between 10 and 20 γ. A similar analysis has enabled us to place limits on the neutral helium content such that the density of neutral helium at an altitude of 100 km must lie in the range of

3×10^7 to 6×10^8 per cm^3, values that are entirely consistent with the amount of helium supplied by radioactive decay from the interior of Venus.

In Figure 3 is shown the thermal structure and ion composition of the topside Venus ionosphere obtained from a self-consistent solution of the heat and momentum transport equations subject to the requirement that the plasma pressure plus the magnetic pressure within the ionosphere, $B^2/8\pi$, balance the solar wind pressure.

The resulting electron temperature, shown in the right-hand portion of Figure 3, is seen to be about 4000° K, with a helium ion temperature, which is the major ion in this particular model, shown to be about 1620° K on the topside, and joining with the neutral temperature toward the bottom. The corresponding neutral gas temperature is about 640° K.

In the left half of Figure 3 is shown the corresponding ion composition. In the lower portion of the ionosphere the dominant ion is carbon dioxide in photochemical equalibrium, with a peak density of 5.2×10^5 per cm^3 at an altitude of about 140 km.

Roughly 100 km above the electron density peak, the scale height changes rapidly in response to the changing ion composition. The large scale height of the topside ionosphere and the narrow electron density peak at 140 km were very prominent features in the Mariner 5 data.

In conclusion, by using a theoretical model to augment measurements obtained by the Mariner 5 spacecraft we have been able to obtain the key parameters that characterize the daytime ionosphere of Venus: the thermal structure, the ion composition, and the magnitude of the horizontal magnetic field.

CHAIRMAN:

Thank you, Dr. Herman. Are there questions?

MEMBER OF THE AUDIENCE:

Is your sensitivity of scale height good enough to tell you which ion is dominant in the topside?

DR. HERMAN:

For a heavy ion such as oxygen, yes, but for the light ion hydrogen, no.

MEMBER OF THE AUDIENCE:

What is the source of hydrogen in your model?

DR. HERMAN:

The source of hydrogen could be several things. There is the hydrogen that comes up from the planetary surface, and there is another possible source that has been suggested by someone who is going to be talking at AGU, Dr. Peter Banks; this is a charge exchange with carbon dioxide. We have not been able, from the abstract he presented, to be sure what the process is, but in my opinion, having worked the process out, I think that it is resonant charge transfer with an excited state of carbon dioxide that is excited about 0.17 eV above the ground state. Thus there is a strong possibility of charge exchange with CO_2 excited plus H^+, and the reverse reaction. This seems enough for H^+ in the Venus ionosphere.

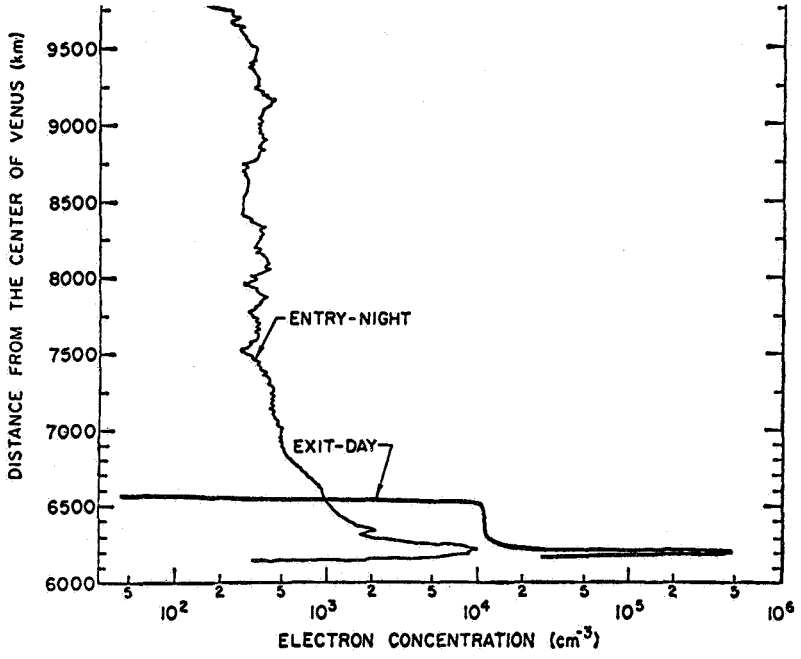


Figure 1

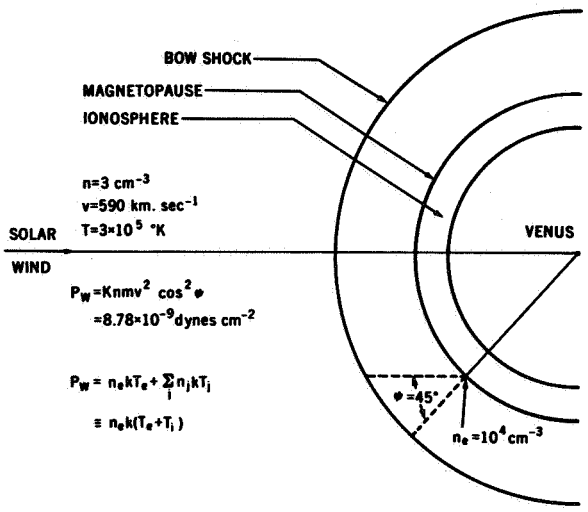


Figure 2

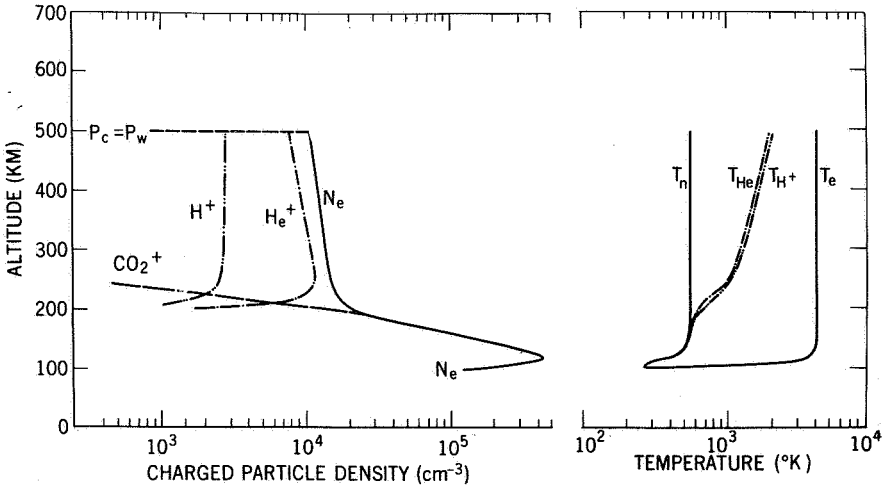


Figure 3

VISCOSITY AND THE SOLAR WIND

John C. Brandt

CHAIRMAN:

The next area with which we deal is solar physics, with a series of papers concerning the nature of the sun and the solar wind. The first of these will be given by Dr. John Brandt.

DR. BRANDT:

In discussing the theory of solar wind, we are in the slightly embarrassing position of finding that the agreement between theory and observation does not necessarily improve when additional and obviously quite needed physics is brought into the problem. It causes a lot of difficulties and requires that we be absolutely sure that advances in physics are included in our theory in a self-consistent manner.

Now, one of the theoretical advances which has led to some difficulty was the realization that the protons and the electrons in the solar wind could not possibly be at the same temperature. These are simply not enough collisions to keep them the same temperature, so one must really consider a two-fluid model of the solar wind in which the proton and electron temperatures are independently determined. This difference in temperatures has subsequently been verified by observations. For quiet solar wind times, times when the wind velocity is somewhere between, say, 250 and 300 km/sec at the earth, the observations indicate that the ratio of electron-to-proton temperature is probably in the range of 5 to 10. This is a difficult measurement to make, but this is what the measurements tell us.

On the other hand, the theoretical models that have been available to this time indicate a rather higher ratio of electron-to-proton temperature, something on the order of 80. So, one might ask the question, what possible physics has been left out, or more correctly, what could we include to somehow raise the temperature of the protons at the orbit of earth?

Now, one way to do this, which has not been included in the theories so far, is to include viscosity, which is a well-known mechanism for taking bulk motions or gradients in bulk motions and converting them into heat or thermal motions. We also know that the coefficient of viscosity varies as the temperature to 5/2-power. This means that viscosity could be important in the solar wind plasma, which starts out at the sun at two million degrees Kelvin and comes down to something like one hundred thousand degrees Kelvin at the earth.

Accordingly, Dr. Charles Wolff, myself, and Bob Southwick, a graduate student from the University of Maryland, began an investigation to look at the possible importance of viscosity in models of the solar wind. The importance of this, I think, is quite apparent and can be seen in Table 1 where the radial momentum equation is given at the top with the inertial, pressure, viscosity, gravity, and magnetic terms labeled. These terms are for a model which we have already calculated. They are tabulated as a function of distance from the sun and are normalized to the pressure term. The sole purpose of this presentation is to show that the viscous terms do in fact appear to be important and should be included, at least in the radial momentum equation. We have also included them in the other equations, such as the energy equation, where required.

This viscous model, when carried through, has a more gentle gradient than the inviscid model and produces a hotter proton temperature at the earth. This can be understood in terms of the velocity gradients near the sun and the fact that if you keep the plasma a little hotter than you had previously, the coefficient of viscosity (proportional to $T^{5/2}$) remains high. You then have a mechanism which feeds upon itself to keep the plasma hotter.

Now this issue is not above debate. The integrations are quite tricky. But on this model, which we have carried through, the ratio of electron-to-proton temperature is about 7. This is essentially in agreement with the observations.

The general conclusion is that viscosity is very likely to be important. We can show that it should be included in the radial momentum equations. When it is included, one obtains a model on which the electron-to-proton temperature ratio for quiet times is closer to the observed values.

DR. HALL:

Is it clear from the work that has been done that the protons and electrons are in equilibrium separately?

DR. BRANDT:

Do you mean that they are essentially isotropic and have a Maxwellian distribution?

DR. HALL:

Yes.

DR. BRANDT:

The answer is no, because of the magnetic field. The velocity distribution is anisotropic with the long axis along the field lines.

DR. HALL:

But is it not near enough so we may use the thermal model idea?

DR. BRANDT:

Yes. Usually the ratio of temperatures parallel and perpendicular to the magnetic field is not over 2 or 3 or something like that, and that value is the closest to unity at the quiet times. So, we do not think that this is a terribly serious error.

Table 1

RADIAL MOMENTUM EQUATION

	$\rho \vec{V} \cdot (\nabla \vec{V})$	∇P	$\nabla \cdot [\frac{2}{3}\mu(\nabla \cdot \vec{V})\vec{\vec{I}} - \mu \vec{\vec{D}}]$	$\rho \nabla \left(\frac{GM_\odot}{r}\right)$	$\frac{\vec{j} \times \vec{B}}{c}$
DISTANCE FROM SUN (a. u.)	INERTIAL	PRESSURE	VISCOSITY	GRAVITY	MAGNETIC
	(NORMALIZED TERMS)				
0.03	0.13	1.00	-0.06	-0.70	0.001
0.30	0.66	1.00	-0.03	-0.24	0.04
0.94	0.94	1.00	-0.01	-0.10	0.05

THE HEATING OF THE SOLAR WIND

Leonard Burlaga

CHAIRMAN:

The next paper concerns the heating of the solar wind and will be given by Dr. Leonard Burlaga.

DR. BURLAGA:

The heating of the solar wind is one of the most important problems in space physics. Dr. Brandt has discussed one aspect of it. Dr. Ogilvie and I have been examining this problem, and in the process we discovered a relationship between the temperature and the bulk speed of the solar wind, which was overlooked by previous workers but which has an important bearing on the heating problem.

The relation is shown in Figure 1. This is a plot of the solar wind temperature, which is proportional to the thermal speed of the solar wind, versus the bulk speed. The observations shown by the open circles are Ogilvie's Explorer 34 plasma data for the period June to December 1967. The basic relation is immediately obvious. The points scatter about the straight line in Figure 1, which is actually a least square fit to the observations. For the period June to December 1967, the state of the solar wind can be represented by a point which is constrained to move up and down this straight line.

The question that next arises is how general the *U-T* relation is and does it apply to other parts of the solar cycle? To answer this, we looked at other data. The point *N* refers to the mean state of the solar wind observed in 1962, one-half solar cycle earlier, by Neugebauer and Snyder on Mariner 2. The point *H* is a summary of data by Hundhauser, data obtained by Vela spacecraft at quiet times during the period from 1963 to 1967. The point *W* is a point for a very quiet time, obtained by Wolfe with the Ames probe in 1966.

You can see that all of these points are very close to our straight line. In other words, the simple *U-T* relationship organizes a great deal of data from all parts of the solar cycle. Thus, we have found a very general and fundamental relationship for solar wind.

Now you may say: So what? What does this have to do with the heating? Well, there are at least three important consequences of this relationship.

First, it puts the existing theoretical work in a new perspective. For example, the two-fluid model of Sturrock and Hartle, which deals with the protons and electrons, is widely believed to be in serious disagreement with the observations. It is thought that the proton temperatures are a factor of 10 too low. But, when you plot the Sturrock-Hartle results on the *U-T* diagram (see the *X*'s in Figure 1), you find that the predicted temperatures are actually consistent with the observations, within a factor of 2, for the very low predicted speeds. This implies then that the two-fluid model, in which heat is supplied only at the base of the solar corona and conducted in the conventional way, is adequate to explain the very quiet state of the solar wind to zeroth order.

The second consequence of the *U-T* relation is that it provides a stimulus for a fresh theoretical approach to the heating problem. Dr. Ogilvie and I suggested, on the basis of some work by Parker, that you may be able to climb up the *U-T* curve by applying heat out to, say, 20 or 30 solar radii. This concept is illustrated at the top of Figure 1 for the extreme case of heating all the way out to infinity in such a way as to get an isothermal corona. In this case, using some work by Parker, one obtains a curve which has, more or less, the right shape, but the temperatures are too high.

Hartle, using a two-fluid model, found that he can reproduce the *U-T* curve up to speeds of about 400 kilometers per second by postulating extended heating out to about 25 or 30 solar radii. So, he can get half-way up the *U-T* curve. The meaning of the curve at higher speeds is still to be explored.

A third consequence is that you can use the *U-T* relationship to study new problems. For example, we used it to identify anomalously hot regions in the solar wind. These regions do not contribute very much to the overall heating of the solar wind, but they do represent regions where an interesting heating mechanism is at work. We do not know what the mechanism is yet but we are working on it, and maybe I can tell you about that next year at this time.

DR. HOFFMAN:

Since the proton temperature is not isotropic, what temperature did you use to organize these data?

DR. BURLAGA:

We used the average temperature. Notice the error bar in Figure 1. That's more or less the scatter you would expect from the observed temperature anisotropy.

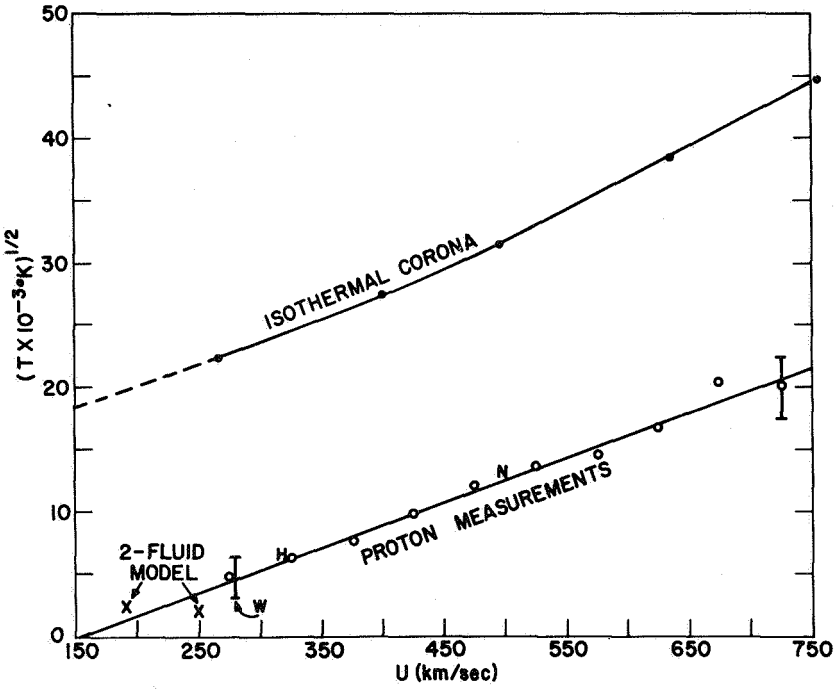

50
40
30
20
10
0
$(T \times 10^{-3} {}^{\circ}K)^{1/2}$
ISOTHERMAL CORONA
PROTON MEASUREMENTS
2-FLUID MODEL
H
W
150
250
350
450
550
650
750
U (km/sec)

Figure 1

THE HEATING OF THE CHROMOSPHERE

Stuart Jordan

CHAIRMAN:

Next we go to the heating of the chromosphere, the paper by Dr. Stuart Jordan.

DR. JORDAN:

It has been conjectured by the French astronomer Cayrel and others that collisional processes can be neglected in determining the population densities and kinetic temperatures in the low solar chromosphere. We know that this is not the case below the low chromosphere, that is, below the temperature minimum in the solar atmosphere, because, in the photosphere, the densities are manifestly too high for collisions to be neglected. And we know that this is not the case above the low chromosphere, because there the theory of spectral line transfer tells us that the emission sources for the strong lines that determine the energy balance in the hotter regions demand the inclusion of collisional processes.

However, the possibility does exist *a priori* that Cayrel's conjecture is correct, and that collisional processes, as, for example, in the case of planetary nebulae, can, in fact, be neglected in the region of the low chromosphere, where the negative hydrogen ion dominates the opacity over much of the spectral range.

Some calculations I have done show the conjecture to be incorrect even in this region, and that, therefore, collisional processes are important in the studies of population densities and in the determination of the kinetic temperature distribution everywhere in the solar atmosphere.

The calculations are fairly easy to perform, and they involve solving only two equations in two unknowns. One of these is the equation for statistical equilibrium for the single-bound level of the negative hydrogen ion, and the other is the energy conservation equation for the problem.

The two unknowns are, first, a so-called departure coefficient (see Figure 1), which we call b_{H^-}, which effectively measures the population density of the negative hydrogen ion, and, second, a temperature, which we call T_e for the electron kinetic temperature, where independent calculations show that it is reasonable to assume that the kinetic temperatures for all chemical species are equal, at least in the chromosphere outside of flares

and other spectacular events. So this is, in effect, the kinetic temperature of the region in question.

To solve the equations it is necessary to evaluate a number of integrals, and, to do this, one must take parameters from various low chromospheric models. To insure that the results I am going to describe are independent of the model chosen, I made the calculations for a large number of models. The general conclusions that I will be describing are independent of the model chosen.

Figure 1 illustrates the results of the calculation. The figure gives three different regions of behavior for the two unknowns, the departure coefficient b_{H^-}, and the kinetic temperature T_e.

The calculation is done as a function of the neutral hydrogen density ranging from values of 10^{11} up to 10^{18} particles per cm^3 and of the effective temperature for the assumed Planckian blackbody radiation field, with this photospheric background radiation temperature ranging from 4,000° K up to 11,000° K.

The top region, called Region I, is the region for which local thermodynamic equilibrium applies. This is the region in which the Saha equation will, in fact, determine the population densities of the various ions and, in particular, the negative hydrogen ion. For this region, the departure coefficient b_{H^-} is equal to 1. The kinetic temperature is simply equal to the minimum temperature, which one obtains from solving the problem after the fashion of the classical stellar atmosphere solution. There is no temperature inversion and no temperature rise; this is because the densities are very high in this region.

Skipping the middle region for a moment, we come down to Region III, sometimes called the planetary nebula case, because here the densities are quite low, generally below 10^{12} particles per cm^3. It is in this region, and this region alone, that the Cayrel conjecture is correct: The departure coefficient and the electron temperature are determined only by the radiation field, and collisional processes can be neglected because of these low densities.

The sun lies right in the middle of the intermediate region where the densities are much too high–way above the planetary nebula case–for collisional processes to be neglected and yet, at the same time, much too low for one to use the Saha equation of thermodynamic equilibrium to determine the population distributions.

Consequently, for the sun, a detailed calculation such as I have made is necessary to determine the values of both the departure coefficient and the kinetic temperature, and these range between limiting values characterizing the very high density region and the very low density region.

The significance of this result, that collisional processes are very important in this region, is the following: If one makes a rough estimate of the upper limit on the amount of mechanical energy required to balance the net radiative losses in this region, one finds that this mechanical energy might exceed by as much as an order of magnitude the amount of mechanical energy required to balance net radiative losses from all other sources of loss of the entire outer solar atmosphere above the temperature minimum.

THE THREE REGIONS OF DIFFERING BEHAVIOR OF (b_{H^-}, T_e) AS FUNCTIONS OF (n_H, T_{eff}). DASHED LINE IS ACTUAL DATA POINT CURVE.

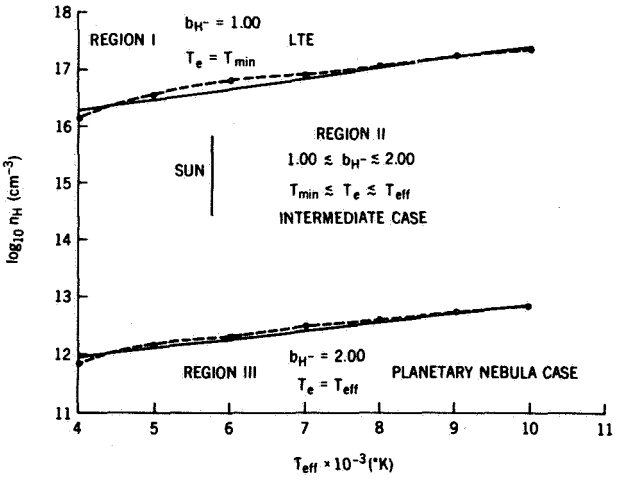


Figure 1

THRESHOLD LAW FOR ELECTRON IMPACT IONIZATION OF ATOMS

Aaron Temkin

CHAIRMAN:

It is necessary, of course, in understanding the sun and other stellar objects, to understand the atomic physics that goes into them. We have at Goddard Space Flight Center a group interested in this effort, headed by Dr. Aaron Temkin. He will talk to us about his work on the threshold law for electron impact ionization of atoms.

DR. TEMKIN:

In Figure 1, I have indicated some of the essentials of the ionization threshold problem, which I will be talking about. The quantity we are interested in is the probability of going from an initial state consisting of an electron impacting on a neutral atom, denoted here by Φ_i, to a final state ψ_f consisting of a positive ion and two receding electrons. The yield Q, then, is the sum over all momenta of the outgoing electrons subject only to the condition of conservation of energy.

Now, if you take away the squared matrix element factor, the remaining integral is the phase space factor, and it can be easily evaluated as being proportional to the excess energy squared.

The threshold law is determined by the limit of the exact wave function ψ_f as the momenta k_1 and k_2 go to zero, the spatial coordinates $\mathbf{r}_1$ and $\mathbf{r}_2$ being finite. This wave function consists of three factors: One factor, ϕ_f, describing the residual ion, which is always considered as a point (positive) charge with no structure; a factor representing the motion of the inner electron; and a factor representing the motion of the outer electron.

The inner electron clearly sees the residual ion; therefore, it is described appropriately by a Coulomb wave, and a Coulomb wave carries the well-known $k^{-1/2}$ factor.

Now the question is, what do we say about the outer electron's wave function?

There are two obvious things one can say. One can say that it, too, sees the charge of the inner core directly, in which case it also is described by a Coulomb wave, and if you then assume it is of this form [(a) in

Figure 1] and plug that right back into the expression for Q, you will find that the threshold law for the yield Q is proportional to the energy excess to the first power.

On the other hand, if you say that the inner electron shields the core from the outer electron, then its wave function is no longer described by a Coulomb wave but by something of shorter range; therefore, it does not carry this $k^{-1/2}$ dependence, and by substitution you will then find that the threshold law is proportional to the three-halves power of E.

On the right of Figure 1, I have given a facsimile of some experimental results. The lower curve is the hypothetical three-halves law, the upper is the linear law. As you might expect, the experimental results are somewhere in between, and it is quite evident that you can never really determine the threshold law from experiment.

Before launching into our approach of trying to determine this effective power, I would like to say that we just heard from Dr. Jordan about nonequilibrium processes; if you can imagine any nonequilibrium process in, say, a solar plasma or any astrophysical plasma, then because it is not in equilibrium, various quantities will be or can be very much dependent on the cross sections involved. Well, ionization is a very good case in point, because the fact of the matter is that if you have a mean temperature much lower than the ionization energy, so that the tail of the electron distribution is very, very low and decreasing very rapidly as a function of the excess energy, then you can see that the whole process (ionization) will be effectively determined by the threshold law itself. I understand that line profiles of the surface of the sun are one such nonequilibrium astrophysical feature and are therefore sensitive to this cross section.

In Figure 2, I have outlined our approach to the problem. On top are two references detailing the approach (as opposed to the results, which are new).

The important point to make here is that there is no rigorous way presently known of solving the many-body Schrödinger equation in the positive energy domain, and this is what makes the problem so very difficult and so challenging.

Our approach has been, therefore, to look in the negative energy domain and try to extrapolate from that. The question that one asks then is, what in the negative energy domain extrapolates to ionization for positive energies?

If you think about it for a while, the answer will come to you that the thing that does extrapolate is doubly excited states, because you can see that if you have two excited electrons, they will both be more likely to come off than if only one were excited.

One can imagine two configurations of doubly excited states: One where both particles are roughly at equal distances from the nucleus, ψ_W, and one where they are not, ψ_D. In the former case it is quite obvious that both electrons will see the nucleus directly, whereas in this asymmetric configuration, the inner electron can shield the outer electron from the nucleus. Physically, I think it is clear that the ψ_W configuration is more likely to lead to ionization than ψ_D.

The point of the calculation then is that since we now believe we understand how to calculate these states as long as we are in the negative energy domain, we can simply compare the energies from these different configurations and, just as in ordinary bound states, decide which configuration is physically likely: the lower the energy, the more appropriate the configuration.

In Figure 3, I give our results; analytically, the point is that the total number of doubly excited states J_D is a linear function of the principal quantum number N. On the other hand, the number of configurations more appropriately described by the symmetrically disposed electrons J_W is given by some dependence N^γ, in which $\gamma \leq 1$. These, giving the major contribution to the threshold law, then effectively diminish the power to which E is raised to $(3\text{-}\gamma)/2$. (A diminution in the power corresponds to an increase in ionization for small E.)

The calculations are being reported here for the first time, and I acknowledge here the very important help of my colleague, Anand Bhatia, in these very difficult calculations. So far, the x's are what we have gotten; the dashed curve represents a least squares fit to the calculations, and it yields a γ of 0.7954, which in turn implies, by the above formula, a threshhold law which goes like E to the 1.102 power.

The results (x's) reported are not final for the following reason: The energies associated with ψ_D for $N>9$ have not been calculated because of the enormous difficulty of that calculation for N large. Rather, we have inferred $J_W(N)$ by counting the number of ψ_W energies that are lower than the $(N-1)$th threshold. We are currently working on a method that will give us the ψ_D energies reliably for large N. Our present inference is based on the

observation that, although the ψ_D energies are approaching the $(N-1)$th threshold energy, they are doing so much more slowly than are the ψ_W energies.

The rather amazing thing about this result is its similarity—practically identity—to a threshold law derived by Wannier in which he described both the electrons and the residual ion as a classical system.

$$Q = \int |\langle \psi_f | V | \Phi_i \rangle|^2 d\underset{\sim}{k}_1 \, d\underset{\sim}{k}_2 \, \delta(E - k_1^2 - k_2^2)$$

$$\text{PHASE SPACE} = \int d\underset{\sim}{k}_1 \, d\underset{\sim}{k}_2 \, \delta(E - k_1^2 - k_2^2) \propto E^2$$

$$\text{THRESHOLD DETERMINED BY LIM} \quad \psi_f = \chi_f(\underset{\sim}{r}_1) \left[\frac{F_c(\underset{\sim}{r}_2)}{k_2^{1/2}} \right] \phi_f(\text{ION})$$

$$k_1, k_2 \to 0$$

$$r_1 \geq r_2 = \text{FINITE}$$

IF OUTER ELECTRON SEES ION

a) DIRECTLY $\Rightarrow \chi_f = k_1^{-1/2} F_c(\underset{\sim}{r}_1)$

$\Rightarrow Q \propto E$

b) SHIELDED BY INNER ELECTRON $\Rightarrow \chi_f = F(\underset{\sim}{r}_1)$

$\Rightarrow Q \propto E^{3/2}$

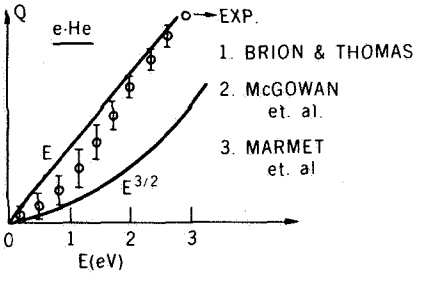


Figure 1—Electron-atom impact ionization.

1. NO RIGOROUS WAY OF CALCULATING ψ_f FOR $E \geq 0$!!
2. OUR APPROACH: CALCULATE FOR $E < 0$ AND EXTRAPOLATE. WHAT EXTRAPOLATES TO IONIZATION? DOUBLY EXCITED (AUTOIONIZATION) STATES!

3. ASSOCIATED WITH EACH N STATE OF TARGET ARE J_D ($\propto N$) D.E. STATES. J_W ($\propto N^\gamma$, $0 \leq \gamma \leq 1$) ARE OF ψ_W TYPE: THE REST ARE OF ψ_D TYPE.
4. J_W (N) CAN BE CALCULATED BY COMPARING ψ_W AND ψ_D ENERGIES!

REFERENCES

TEMKIN, BHATIA, SULLIVAN
PHYS. REV. 176.80 (1968)

A. TEMKIN, IN INVITED PAPERS OF THE SOMMERFELD CENTENNIAL MEMORIAL CONF. (NORTH-HOLLAND, 1969).

Figure 2–Doubly excited state approach.

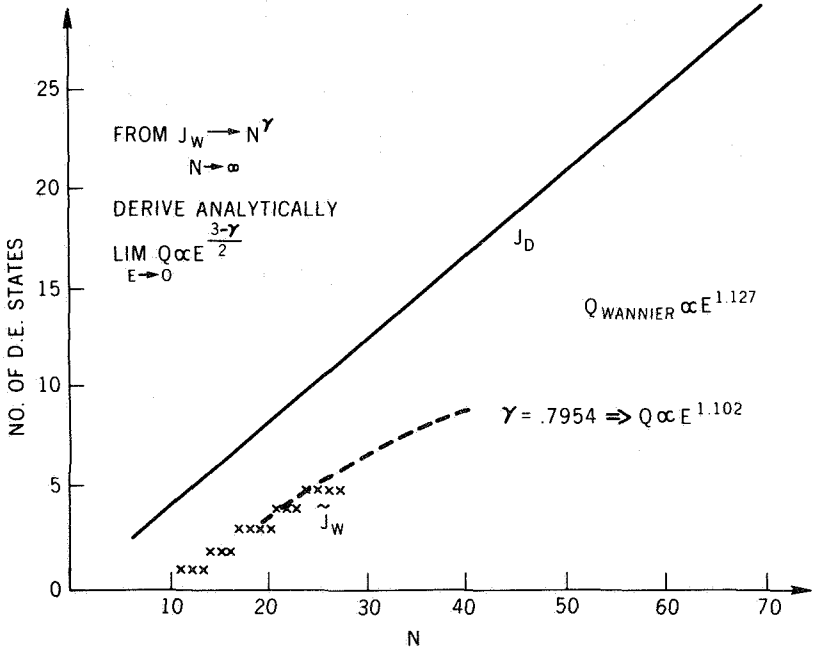


Figure 3–Results (1969).

OBSERVATIONS OF THE ULTRAVIOLET SOLAR FLUX VARIABILITY

Donald F. Heath

CHAIRMAN:

We turn now to observations of the Sun and solar emissions of various types. The first paper, by Dr. Donald Heath, concerns observations of the ultraviolet solar flux variability.

DR. HEATH:

On April 14, 1969, an experiment designed to monitor the Sun in the meteorologically important ultraviolet region was launched aboard NIMBUS 3 into a circular, near-polar orbit. Since then, three vacuum photodiodes have monitored the solar radiation in the 1300- to 3000-Å region for about 20 minutes of each 107-minute orbit. These observations also include observations of atmospheric occultation of the Sun.

The data I am presenting are those recorded at the northern terminator. Up to ten crossings per day are recorded and are then averaged together to give one point per day, as shown in Figure 1. (Readings from a particular sensor may agree to within one telemetry bit in 128.) The daily averages of the sensor outputs are shown as the ordinate, the day number as the lower abscissa, and the solar rotation number as the upper abscissa.

Radiation detected by the 2750- to 3150-Å sensor comes from the photosphere, that in the 1750- to 1850-Å range comes from the top of the photosphere, and the 1350- to 1600-Å radiation originates at the base of the chromosphere. For comparison, the daily averages of the 10.7-cm solar flux are shown at the bottom of the figure.

The 1350- to 1600-Å radiation is observed with a diode consisting of a magnesium fluoride window and a solid tungsten cathode. Initially 40% of this signal was due to Lyman alpha. However, radiation damage produces an absorption band that blocks Lyman alpha. This is verified by observations of atmospheric occultation of the Sun.

Since launch, the signal intensities have decreased approximately by a factor of 6. Two effects are readily apparent: There is a long-term signal decrease as well as an effect that is periodic with the solar rotation number. A graph of the logarithm of the signal versus time should exhibit either a single straight line or multiple straight lines decaying with time. For this type of decay one would expect the values of the sensor readings to decrease with time. If one were to examine such a curve, he would

observe a maximum (sensor reading) value in May, a minimum in August, and a maximum around November. This variation in ultraviolet radiation may contribute to the semi-annual variation of atmospheric density by altering the dissociation rate of O_2 at 120 km.

The long-term variations of the 2750- to 3150-Å sensors are 20%, and the long-term variations of the 1750- to 1850-Å sensor are only 6%.

Figure 2 shows the signals on an expanded scale after corrections for the changing Earth-Sun distance have been applied and an exponential decay factor has been removed. The 1750-Å and the 1350-Å radiation regions are separated by what is known as the solar temperature minimum. The radiation from the top of the photospheric region corresponds to a temperature of 4650° K. The maxima and minima in the readings of the 1750- to 1850-Å and of the 1350- to 1600-Å sensors correlate extremely well with the maxima and minima of the 10.7-cm flux. The basic difference is that the radiation coming from the top of the photosphere has an essentially constant variation with rotation of about 6%, whereas the radiation from the base of the chromosphere has variations of as much as 47%, although at most times the variations are much less.

The observations are continuing, and one feature I would like to point out is that the sizes of these large maxima have no correlation with the sizes of the maxima of the 10.7-cm flux. In fact, it turns out that these observations show more detail than can be seen in the variations of the 10.7-cm flux.

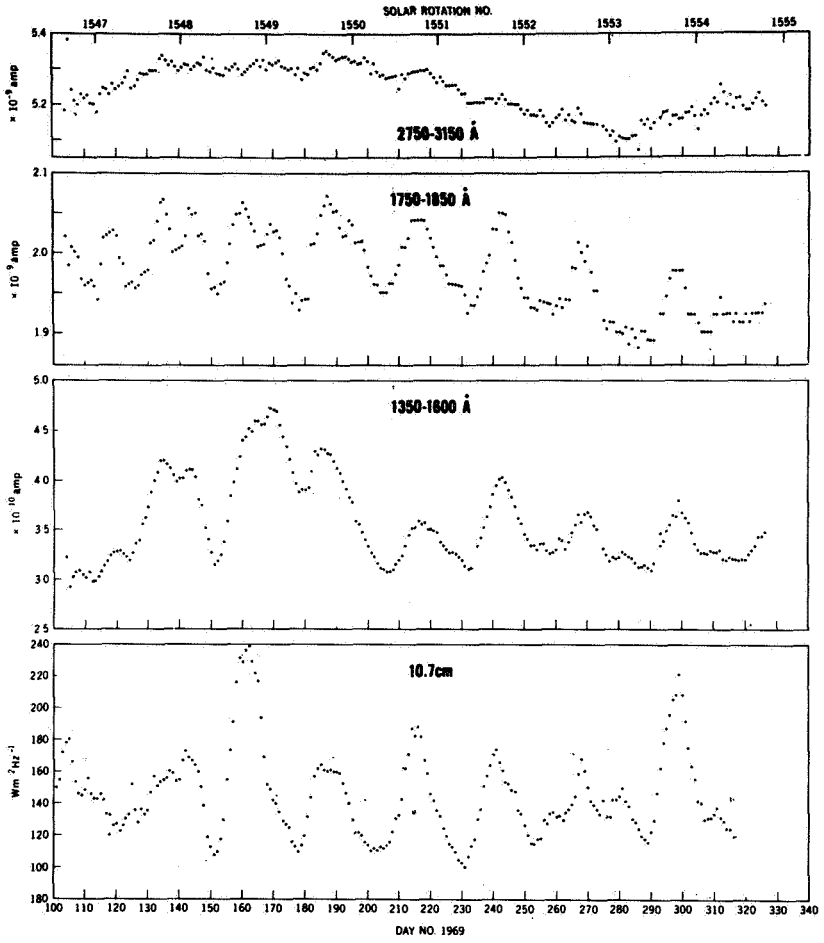

SOLAR ROTATION NO.
1547
1548
1549
1550
1551
1552
1553
1554
1555
2750-3150 Å
1750-1850 Å
1350-1600 Å
10.7cm
DAY NO. 1969

Figure 1

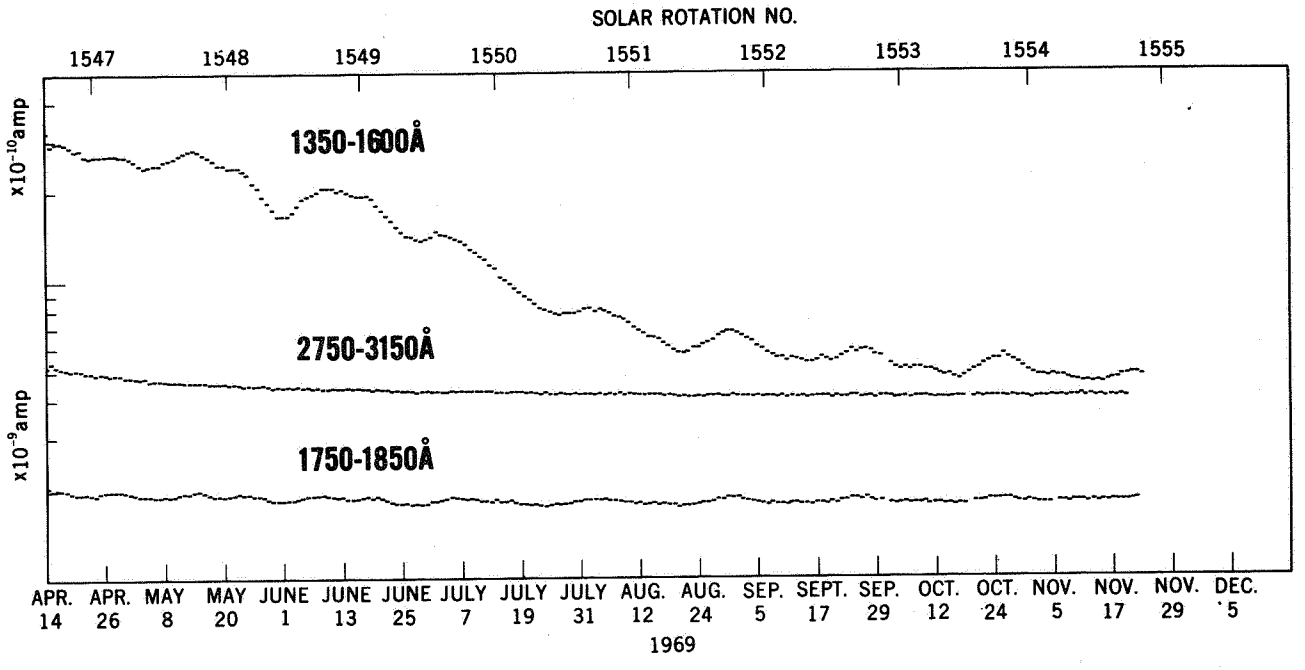

SOLAR ROTATION NO.
1547
1548
1549
1550
1551
1552
1553
1554
1555
x10⁻¹⁰amp
1350-1600Å
2750-3150Å
x10⁻⁹amp
1750-1850Å
APR. 14
APR. 26
MAY 8
MAY 20
JUNE 1
JUNE 13
JUNE 25
JULY 7
JULY 19
JULY 31
AUG. 12
AUG. 24
SEP. 5
SEPT. 17
SEP. 29
OCT. 12
OCT. 24
NOV. 5
NOV. 17
NOV. 29
DEC. 5
1969

Figure 2

SOFT SOLAR X-RAYS

Werner M. Neupert

CHAIRMAN:

Next, we move to the region of soft solar X-rays and hear from Dr. Werner Neupert.

DR. NEUPERT:

I am going to discuss information on solar flares that we derived from solar X-ray spectra between 1 and 8 Å. These spectra were obtained by OSO 5, launched in January 1969. We had already found from our OSO 3 observations in 1967 that this region of the spectrum represented thermal emission from the hot plasma that evolved during the initial stages of the flare event. Our observations from OSO 5 have improved to the point where we can now begin to get information on detailed physical properties of this flare plasma.

Figure 1 shows the solar spectrum associated with two very different levels of solar activity, with the counting rate of the detectors plotted against wavelength. The lower of the two spectral distributions is the solar spectrum in the absence of flare activity. The scale for this spectrum has actually been expanded by a factor of 5 so that we may see it better.

There are two components in the lower spectrum. The first is a continuum which is characteristic of radiation from an optically thin plasma at a temperature of about six to eight million degrees Kelvin. Superimposed on this continuum is line emission. In this case the emission is from helium-like silicon XIII.

The upper curve is an observation recorded during a solar flare of importance 2B on February 27, 1969. Notice that the spectral distribution has changed markedly. First, the continuum distribution has moved toward shorter wavelengths and is now characteristic of a plasma at about 27 million degrees Kelvin. Second, we observe many more emission lines on this continuum. They are all from very highly ionized elements up to nickel in the periodic table. For instance, we see Fe XXVI, which is iron with only one electron left. Again, the temperature in the plasma required for such stages of ionization is of the order of 20 to 30 million degrees Kelvin.

Figure 2 shows with greater clarity a small portion of the flare spectrum of Figure 1. In the region around 6 Å, we observe the helium-like

emission from Ca XIX and Si XIII. Notice that there are three prominent lines associated with each ion. The right-hand line in each case has not been observed in laboratory plasmas and was not expected to be in the solar spectrum. It is actually a forbidden transition. The interesting point is that the intensity of this forbidden transition depends upon the electron density in the region where the emission takes place. From the observed intensity of the forbidden transitions, we infer that the density in the flare region is of the order of 10^{11} electrons/cm^3, much greater than in the normal corona but lower than that encountered in laboratory plasmas.

The information that we get from the three lines emitted by a helium-like ion such as Si XIII is the electron density, obtained from the intensity of the transition relative to the other lines, and the total amount of material that is emitting at the temperature at which the ion exists. The latter is obtained from the total emission in all three lines. Given these two facts we can infer the volume from which the emission takes place, and this volume is very small on the scale of solar dimensions. It may be, for instance, a filament about 50,000 km long having a diameter of 1000 km.

No X-ray photograph has such resolution, but we anticipate that, with the larger instruments now being developed for the Apollo telescope mount and those planned for the future, we will be able to see the actual and very interesting structure of such extremely hot regions on the Sun.

CHAIRMAN:

Is this the first time that one has been able to determine the volume of a flare-emitting region?

DR. NEUPERT:

This is certainly the first time. I think our observations indicate that the volume is smaller than we expected. The interesting point about it is that we are pushing the theoreticians to refine their models of the flare mechanisms.

MEMBER OF THE AUDIENCE:

Could you say something about the energy in the flare?

DR. NEUPERT:

The total energy in the flare volume, for a large flare, is of the order of 10^{32} ergs. Only a part of that is in the form of radiated energy. Of

course, these flares usually have very high kinetic energy associated with the ejection of material and high energy particles.

MEMBER OF THE AUDIENCE:

What was the source of the energy in a flare?

DR. NEUPERT:

We really do not know. It has been suggested that annihilation of magnetic fields may be the source. The difficulty is that no one has made what can be agreed to be a definitive observation of the actual change in magnetic fields during the event. That remains to be made. Another process suggested is that in which high-energy electrons or protons are stored in the corona prior to the flare and, by some mechanism that we do not understand yet, are precipitated into the chromosphere, producing an observable effect. Again, there are problems with how the particles are stored in the first place without their energy dissipating. The question is, how is this amount of energy stored?

MEMBER OF THE AUDIENCE:

What is the basis for assuming that the plasma is in thermal equilibrium?

DR. NEUPERT:

It is probably that the relaxation time, the time required for all the electrons to achieve a Maxwellian distribution, is really of the order of 10 to 100 seconds at the electron densities we are talking about, which are of the order of at least 10^{10} electrons/cm^3 and perhaps as high as 10^{12} electrons/cm^3. This time is comparable to or shorter than the duration of time over which we see a flare event. The flare spectrum shown in Figure 1 and Figure 2 was maintained for 10 minutes; that is long compared with typical relaxation times and therefore suggests to us that we are looking at a thermal phenomenon.

MEMBER OF THE AUDIENCE:

Is it possible that the union of light nuclei to form the heavy ions that you are talking about is incomplete?

DR. NEUPERT:

That is a field I am not familiar with. It is true that one does not really see the production of high stages of ionization in a flare. It is not as

if we start with a cool plasma and heat it up more and more. It just seems to appear; we always see the plasma cooling off, but we never observe it heating up.

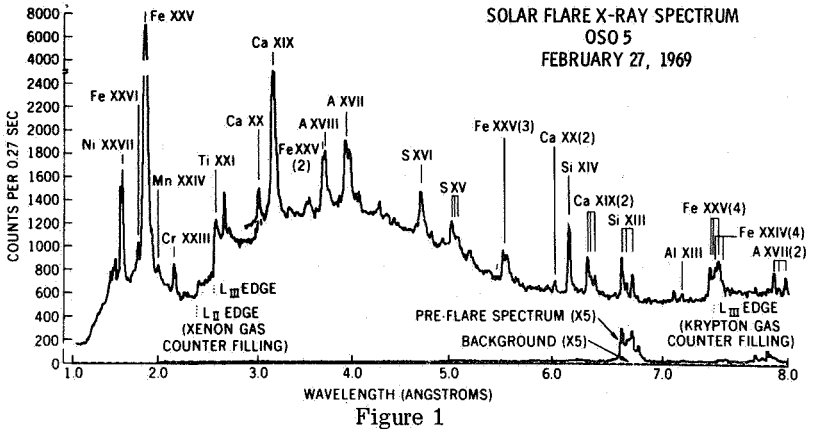


Figure 1

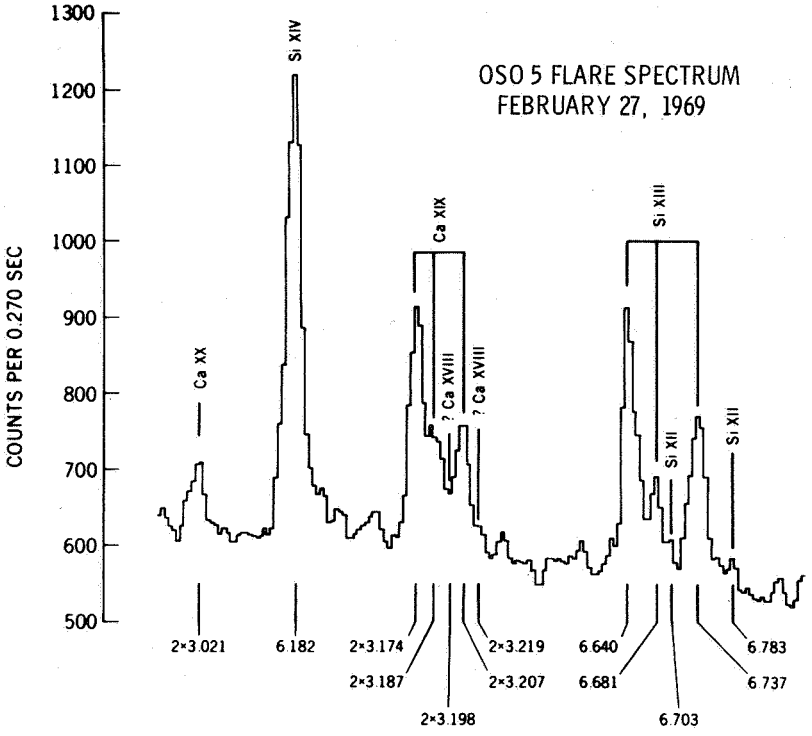


Figure 2

ENERGETIC SOLAR X-RAYS

Kenneth J. Frost

CHAIRMAN:

Next, Kenneth Frost will discuss energetic solar X-rays.

MR. FROST:

I would like to talk to you about the OSO 5 measurements of X-rays in the energy range from 14 to 250 keV. These were made with a time resolution of 1 sec over the entire energy range which had been subdivided into nine energy intervals. X-rays in this range are called hard solar X-rays to differentiate them from the softer photons discussed by Dr. Neupert.

Hard X-rays are produced by free electrons moving with velocities up to 0.8*c*. When these electrons are slowed down or retarded in materials whose densities are typical of those encountered in the solar chromosphere, a fraction of the energy lost by the electron beam appears as hard X-rays. This process of X-ray generation is known as bremsstrahlung, from the German word meaning braking radiation.

Radiation in this range is also typical of the solar flare since the Sun is undetectable in the 14- to 250-keV energy interval when it is not flaring.

In Figure 1, we have the data obtained for an event on March 1, 1969. This X-ray flare was associated with an explosive prominence of importance 3 on the west limb of the Sun. On the left side of the figure, the X-ray burst is shown in three of the channels: 14 to 28 keV, 28 to 55 keV, and 55 to 82 keV. The higher energy channels are not plotted, but they have about the same intensity-time profile except at reduced intensity. The right side of the figure shows the correlated microwave radio data at 2,800, 10,700, and 15,400 MHz.

An interesting feature to note in this X-ray burst is the change in its character from low to high energy. In the lowest energy channel, the broad emission peak dominates, with only one or possibly two initial sharp spikes evident. In the higher energy channels, the broad emission feature disappears, and the initial spikes dominate. This energy dependence suggests that the spikes are nonthermal while the broad peak, occurring later in the flare, is at least partly thermal in origin.

This observation is quite typical of a large class of events we have now observed with OSO 5; the X-ray burst begins with a repetitive and, in this case, quasi-periodic group of fast X-ray bursts having a very hard spectrum. This repetitive X-ray emission is then followed by a slow-moving, thermal emission.

Figure 2 shows the photon spectra associated with each of the features in Figure 1. The dots represent the spectrum observed for each of the three fast spikes in the beginning of the flare, and the triangles show the spectrum of the later broad peak. The most interesting characteristic here is that the photon spectrum of the spikes can be fit with a power law, $E^{-3.2}$, up to about 100 to 150 keV. At this point, there is an abrupt change in the slope of the spectrum indicating that the accelerating mechanism did not produce electrons with energies above roughly 300 keV.

The photon spectrum observed for the broad maximum is fit by a power law, E^{-5}, over the entire detectable energy range. However, this spectrum does not arise from a single source, but is probably composed of a waning, nonthermal radiation from the first three spikes and the rising thermal radiation which dominates later in the event. Near the end of the event the spectrum can be fit quite nicely by a thermal $\exp(E/kT)$ with a temperature of the order of $10^8\,^\circ$K.

The final point I wish to make is that it has been assumed or explicitly stated by some workers that the thermal emission observed later in the event arises from the thermalization of the initial fast electrons. We do not believe that this is the case. The picture we have in mind is that, during the compression stage of the plasma, groups of electrons are shot out repetitively and sometimes periodically spaced with velocities up to 0.8c, and that these electrons are responsible for the hard X-ray emission we see initially. The thermal emission follows later as the compressed plasma begins to cool.

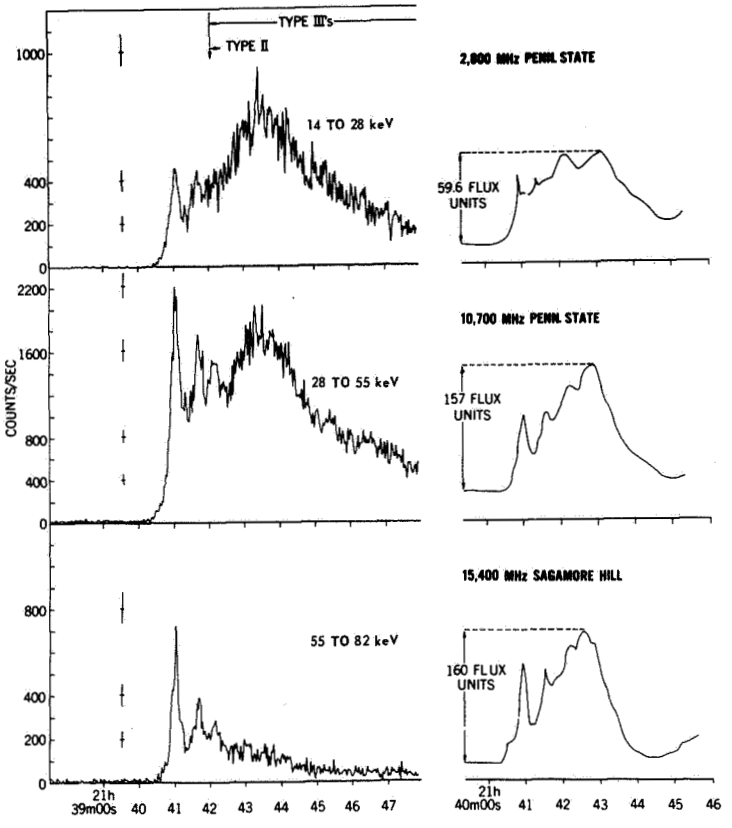


Figure 1

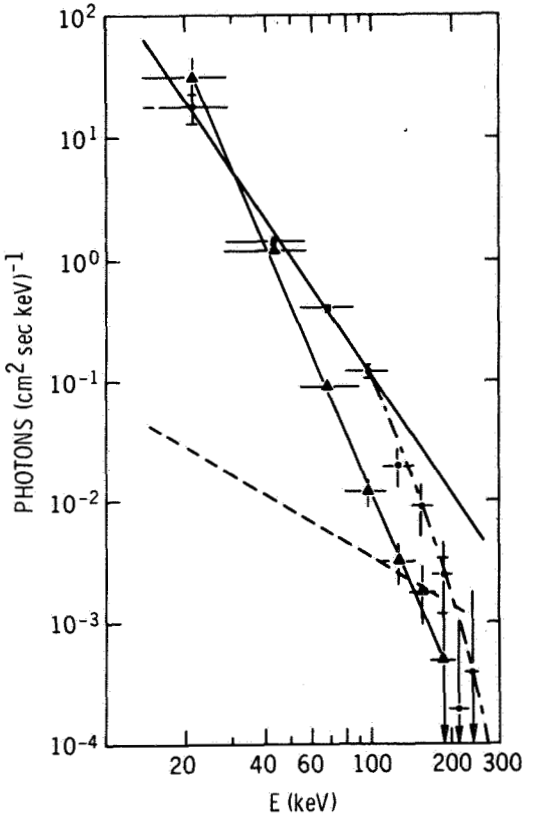


Figure 2

SOLAR RADIO EMISSIONS

Joseph Fainberg

CHAIRMAN:

We turn now to the subject of solar radio emissions and the paper by Dr. Joseph Fainberg.

DR. FAINBERG:

In a 2-week period of August 1968, a storm of tens of thousands of rapidly drifting radio bursts was observed by the first Radio Astronomy Explorer (RAE) satellite. Although a non-directional dipole antenna was used for these observations, we nevertheless have identified the active region of the sun from several possibilities.

We have derived from the experimental data an average exciter velocity of the bursts and an average density profile in the corona above the active region, and we have obtained some measure of the Archimedes spiral curvature of the streamer itself in the region of 11 to 40 solar radii. This involved a new type of analysis based on the systematic change of drift times with solar rotation. These radio bursts are the low frequency counterparts to the rapidly drifting type III bursts observed at higher frequency by ground-based instruments.

The bursts are believed to originate from short packets of energetic electrons traveling outward through the corona and exciting progressively lower frequencies. The ground-based observations are consistent with coronal streamers of enhanced density existing out to about 5 solar radii above active regions. Now, to extend these results to greater distances out into the corona requires radio observations at considerably lower frequencies. This is possible only from above most of the earth's ionosphere.

Prior to RAE 1, space observations indicated the existence of radio bursts down to quite low frequencies. However, because of the lack of directional information, it was not possible to derive from the data either values of exciter velocity or information about the curvature of the streamer.

Figure 1 illustrates the geometry of the analysis.

As the exciter moves outward from the sun, it excites plasma oscillations of progressively lower frequencies. The drift time of the burst from the higher frequency f_1 to the lower frequency f_2 is determined primarily by

the time it takes the exciter to travel the distance of the level separation ΔR. However, at the earth this drift time is reduced by the fact that the travel path from the lower frequency is shorter than the travel path from the higher frequency. This observed drift time will be at a minimum when the emission region crosses the sun-earth line.

Note that, because of the curvature of the streamer, this time of central meridian passage will occur later for radio emissions further out along the streamer. As the streamer rotates from the east limb of the sun to the central meridian, we expect to see a progressive decrease in drift time followed by an increase as the streamer rotates around to the west limb.

In Figure 2 we have given a portion of the experimental data involving several thousand measured bursts. Plotted here on a daily basis for a 2-week period in August 1968 is the histogram of the number of bursts plotted vertically versus the drift time plotted horizontally in seconds.

Note the progressive decrease in the average drift time from August 12th to about August 20th, followed by a progressive increase through the 28th and 29th. This we associate with the rotation of the streamer from the east limb to the west limb. The dots on this figure are the result of the least-squares fit of the experimental data to the best derived values of exciter velocity, of level separation, and of time of central meridian passage.

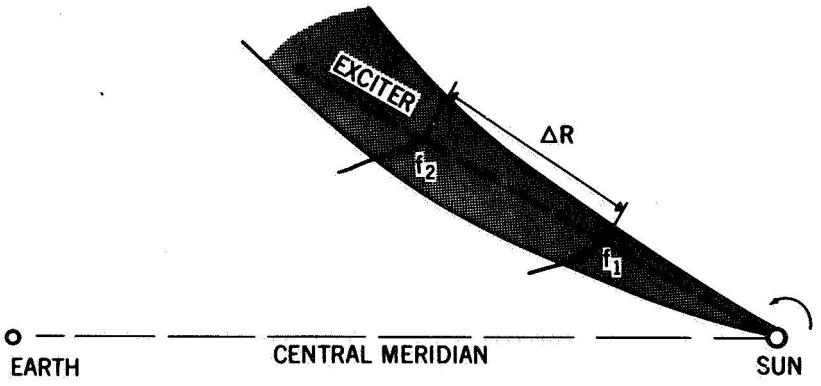
This group of data is for a particular pair of frequencies. We have used six separate frequencies in this analysis, so we have a similar group of data for the other possible pairs of frequencies.

Figure 3 illustrates some of the results of the analysis.

First is the existence of a streamer of enhanced density in the region of 11 to 40 solar radii. The average density enhancement has been measured to be 15 times that of the quiet sun at solar minimum.

The average exciter velocity has been found to be 0.36 times the velocity of light. This is quite consistent with ground-based observations at higher frequencies and lower solar elevations, and it is also consistent with short groups of 40-keV electrons observed by Lin and Anderson in the vicinity of the earth.

A third result is the progressive increase in central meridian passage time with radial distance outward along the streamer. From values that we have measured, we can derive an average bulk velocity of the material in the streamer between 11 and 40 solar radii of 370 km/sec.



DRIFT TIME = f(LEVEL SEPARATION, EXCITER VELOCITY,
TIME OF CENTRAL MERIDIAN PASSAGE, TIME)

Figure 1 – Streamer geometry.

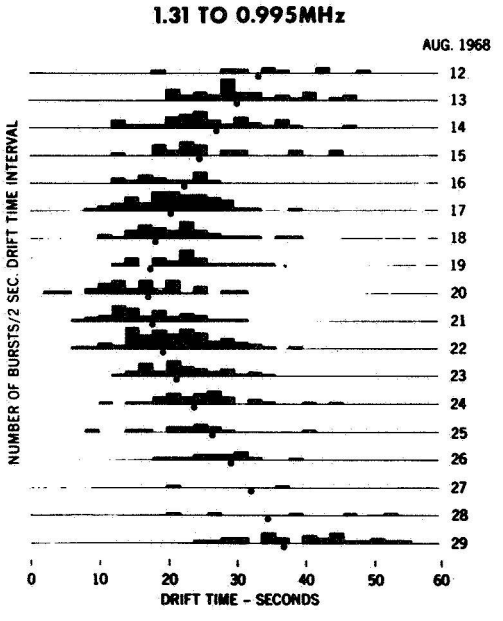
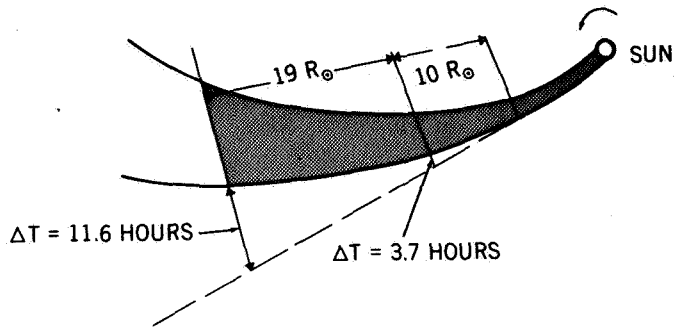


Figure 2



EXISTENCE OF STREAMER	11 TO 40 $R_\odot$
AVERAGE DENSITY ENHANCEMENT	15 x (QUIET SUN AT SOLAR MINIMUM)
AVERAGE EXCITER VELOCITY	0.36 c
AVERAGE BULK VELOCITY	370 km/s

Figure 3 – RAE 1 results.

RELATIVISTIC INTERPLANETARY ELECTRONS AND POSITRONS

Thomas L. Cline

CHAIRMAN:

Next we will hear from Dr. Thomas Cline on the subject of relativistic interplanetary electrons and positrons.

DR. CLINE:

Last year, we introduced the study of the interplanetary component of electrons with energies specifically in the region of a few million electron volts. These particles are interesting because they have the lowest kinetic energies of all particles that are candidates for cosmic radiation, yet, since they are relativistic, they have the same ratio of energy to mass as do most cosmic ray nucleons. We described the quiet-time (and possibly galactic) component of the interplanetary electrons and the separately identifiable solar-flare component, both of which were found in GSFC investigations.

In the intervening year, several developments have occurred that have a direct bearing on the question of the cosmic-ray nature of the quiet-time electrons; they also raise the question of the existence of a solar component present at times other than during the obvious flare or storm occurrences. These developments stem from our investigations of the properties of the interplanetary electron population. These properties include the temporal and spatial behavior of the electrons and the algebraic charge distribution. In principle, these properties may ultimately allow us to trace the various origins of the interplanetary electron population.

We are studying the positron-to-electron ratio in several energy regions. Figure 1 shows the intensities of interplanetary positrons having energies above 2 MeV, observed by several groups. A best fit to a composite of the intensities of the total component of electrons and positrons is also shown. Since data for the electrons have much smaller relative errors, we can plot the various observations in this way simply to indicate the drift of the data. All of the various satellite and balloon positron measurements are shown with the same symbols, and our results from OGO 5 are those at the lowest energies (below 10 MeV).

Comparison between the electron and positron intensities shows that, in the energy interval of 2 to 10 MeV, positrons are only a few percent as numerous as electrons. (Because the electron spectrum is quite a bit

steeper at these low energies, the positron fraction is about one-tenth of the electron value at medium energies, although this may not be obvious to the eye.)

The positron fraction is so small, in fact, that, since a negligible proportion of the electrons is from the pion-decay source (assumed to be an important effect in the 100-MeV region), we can conclude that the 2- to 10-MeV electrons indeed form a separate component in the interplanetary electron population. Modulation, except for that caused by electric fields, would be charge independent. Therefore, we can assume that this component has a separate origin and is not simply a kink in the spectrum caused by an energy-dependent solar modulation.

To investigate the question of how many of these 2- to 10-MeV electrons (which are essentially all negative) are of solar origin and how many remain of galactic origin, we turned to the time variations of the intensity. The daily intensity history for the total electron component over the last 2 years is shown in Figure 2. Here we see a maze of intensity variations, of different sizes and shapes, that were observed with Frank McDonald's cosmic-ray apparatus which was aboard several recent IMP satellites. This apparent chaos can be reduced so as to show the presence of four coherent populations.

First, there are the solar flare increases indicated by short vertical bars. Of course, most of the solar flare electron events are so intense as to be off the graph. Those increases that do not go off the graph can be unambiguously identified because they are coincident with familiar phenomena, such as microwave and X-ray bursts. Second, we have an identifiable component of storm particle or recurrence events. Third, there is an irreducible minimum intensity; one question is how much of it is of galactic origin. Fourth, there are slow time variations of an anomalous nature. For example, in late 1967, four or five of them are obvious; their maxima and minima are identified by shaded areas and arrows, respectively, in the figure.

Thus, in addition to the flare particles, storm particles, and the genuinely quiet-time electrons, there is an electron component occurring at what would otherwise be termed quiet times. These electron intensity variations have an interesting inverse correlation with the proton intensity. For comparison, Figure 3 shows the very-low-energy interplanetary proton intensity. The shaded areas, which correspond to maxima in electron intensity, always occur at proton intensity minima; the arrows referring to electron minima always occur at proton maxima.

This apparent inverse correlation seems to be a coherent phenomenon or property of the quiet-time electron population. We noticed the first example of this effect with the IMP 1 data in 1964 in a series of electron increases spaced 27 days apart, 2 weeks out of phase with proton increases. The inverse correlation might then be dismissed as only a possible randomly phased series. Now, however, with results halfway through a solar cycle, we seem to be observing a fundamental, or at least a very persistent, property of the quiet-time interplanetary electron population. It is as though the Sun emitted electrons and protons in oppositely directed beams.

This feature was unexpected and is unexplained at the present time. Ultimately, it may be identified with a solar modulation of an intrinsically quiet-time interplanetary electron population or with a separate emission of electrons by the Sun. At least its isolation is consistent with the hypothesis of a residual galactic electron component.

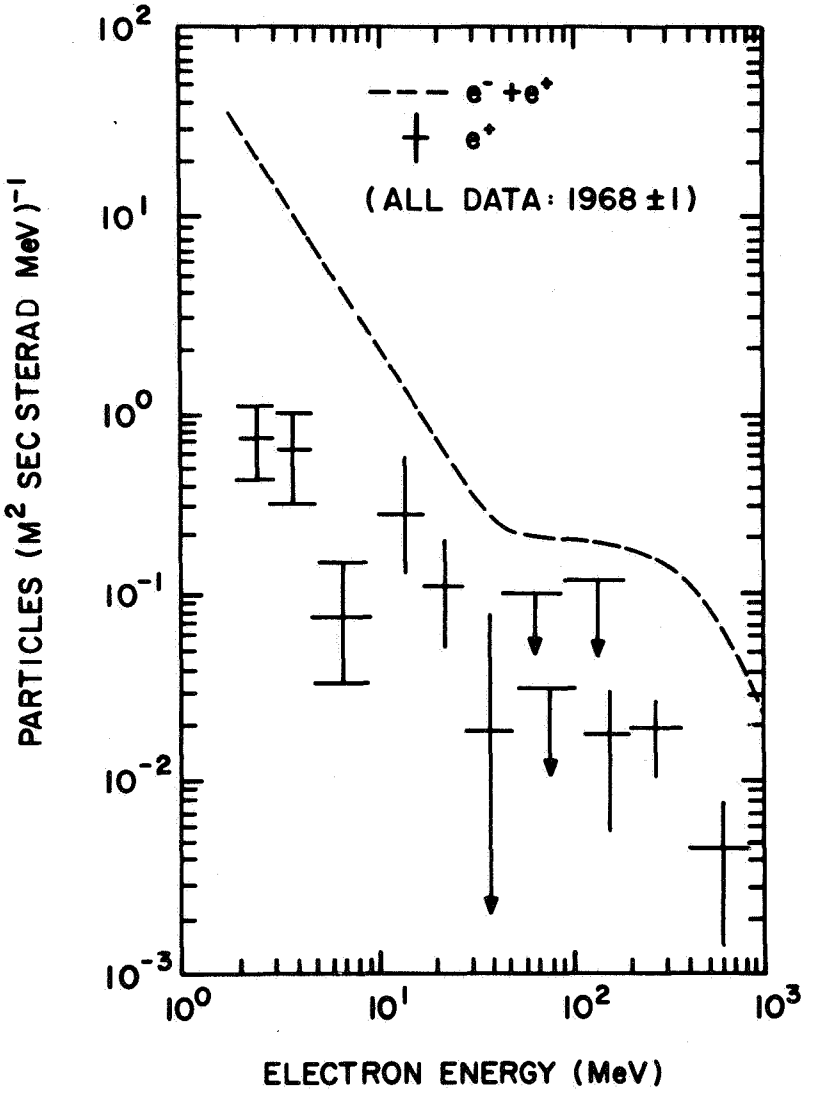

10^2
10^1
10^0
10^{-1}
10^{-2}
10^{-3}
10^0
10^1
10^2
10^3
$e^- + e^+$
e^+
(ALL DATA: 1968 ±1)
PARTICLES (M^2 SEC STERAD MeV)$^{-1}$
ELECTRON ENERGY (MeV)

Figure 1

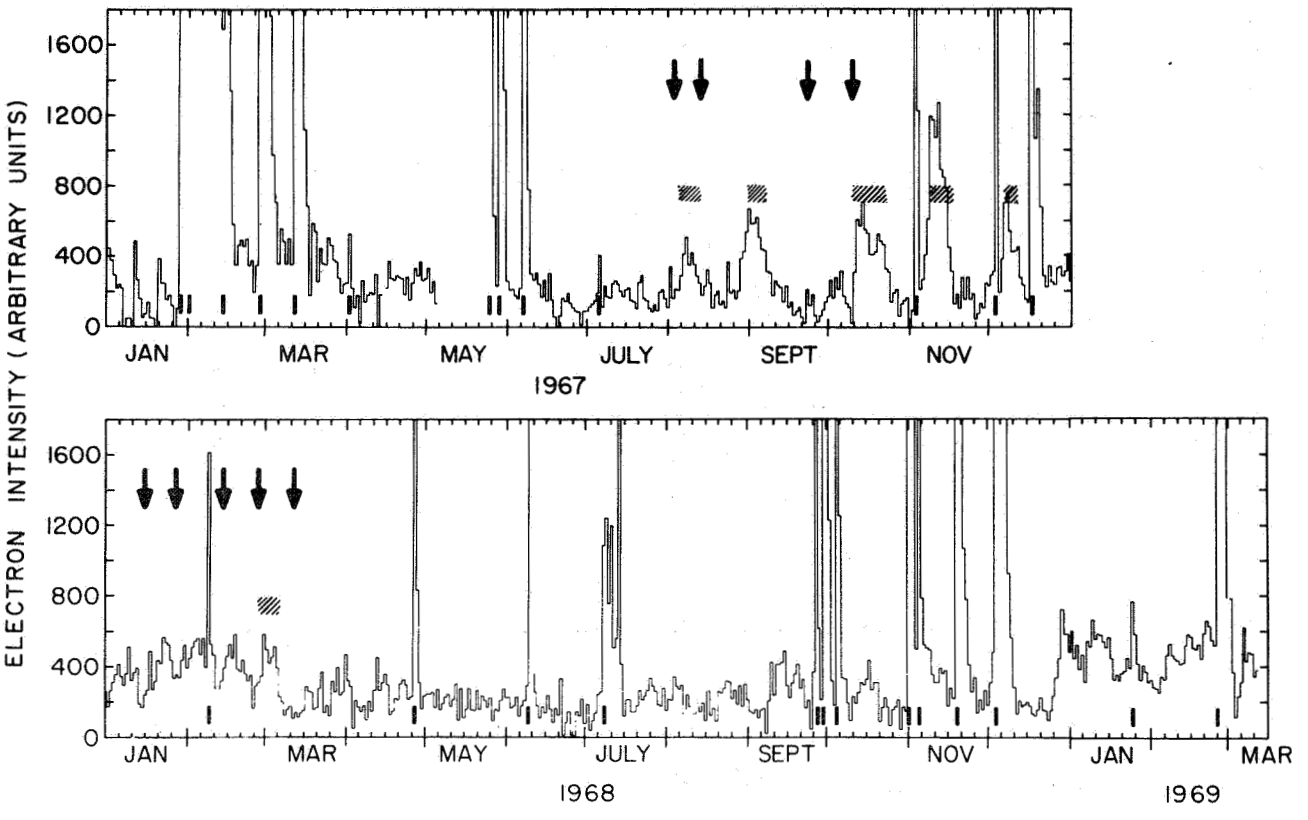

ELECTRON INTENSITY (ARBITRARY UNITS)
1600
1200
800
400
0
JAN
MAR
MAY
JULY
SEPT
NOV
1967
JAN
MAR
MAY
JULY
SEPT
NOV
JAN
MAR
1968
1969

Figure 2

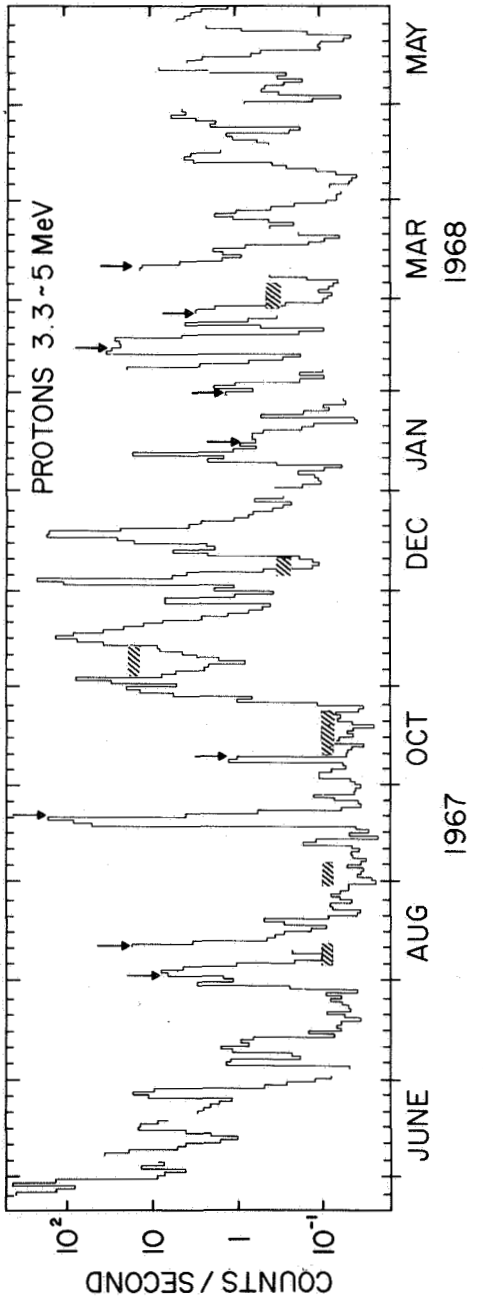

PROTONS 3.3-5 MeV
COUNTS / SECOND
10^2
10
1
10^{-1}
JUNE
AUG
OCT
DEC
JAN
MAR
MAY
1967
1968

Figure 3

GALACTIC RADIO EMISSIONS

Joseph K. Alexander, Jr.

CHAIRMAN:

The next paper is concerned with galactic radio emissions and will be given by Mr. Joseph Alexander.

MR. ALEXANDER:

One of the primary objectives of the Radio Astronomy Explorer satellite was to see what new information we could get about the nature of the galaxy in the neighborhood of the Sun by extending our radio observations down to very low frequencies. I am going to describe today how we use measurements of the average spectrum of the background radiation from the Milky Way to estimate the density of thermal electrons in the interstellar medium and to make some estimates of the flux of interstellar cosmic rays at low energies. Dr. Cline has already shown that it is difficult to measure these quantities at the Earth because of solar modulation effects.

In Figure 1, a plot of the specific intensity or average brightness as a function of frequency shows a sample of the average background radio spectrum over the frequency range of a few hundred kHz to a hundred MHz. The data points from below 1 MHz up to about 9 MHz are from the satellite measurements. At the higher frequencies, from about 9 MHz to 100 MHz, the measurements are made with scaled antennas located on the ground. All of the measurements refer to relatively large areas over the sky and are, in this case, averages of the whole sky brightness.

The power-law increase in the intensity as frequency decreases in the high frequency region is indicative of synchrotron radiation, presumably from cosmic ray electrons. The maximum in the spectrum and subsequent decrease in intensity at the lowest frequencies are an indication of absorption of the synchrotron radiation by the ionized component of the interstellar gas.

From analysis of the radiation spectrum and the way it is absorbed by the interstellar gas, we can make some estimates of the properties of the interstellar medium. At the lowest frequencies, below 1 MHz, all the radiation we are seeing in the phenomena that we are able to examine is occurring near the Sun in the neighborhood of the local spiral arm. To show you how we get some information about the nature of the spiral arm, let me very briefly describe our procedure.

The observed intensity I is a function of the synchrotron radiation intensity I_0, modified by a function of the optical depth τ, which describes how the radiation is absorbed by the thermal electrons at low frequencies:

$$I = I_0 (1 - e^{-\tau})/\tau,$$

where

$$\tau \simeq \frac{\text{CONST.}}{f^2} \frac{N_e^2 L}{T_e^{3/2}} .$$

From an analysis of the way the spectrum falls off at the lowest frequencies we can deduce τ, which is a function of the interstellar electron density N_e and interstellar electron temperature T_e over the path length L that we are sampling—the path length over which the radiation is absorbed. From an analysis of the spectrum, such as the one mentioned previously, an estimate of the interstellar thermal-electron density of the order of .03 electrons/cm^3 is obtained, which is consistent with current thoughts about the nature of the interstellar medium, measurements from pulsar dispersion, and other techniques of that sort.

The synchrotron radiation intensity I_0 is a function of the path length L over which the radiation occurs and of the galactic magnetic field H in the source region:

$$I_0 = (\text{CONST.})\, LKH^{(\gamma+1)/2} f^{-(\gamma-1)/2}$$

where γ and the proportionality constant K are determined by the cosmic ray electron energy spectrum. We assume that the flux of cosmic ray electrons is proportional to the energy of the electrons raised to the power $-\gamma$:

$$N(E) = KE^{-\gamma}.$$

So from an analysis of the radiation spectrum at the lowest frequencies, we can get an estimate of the energy spectrum of the low-energy cosmic ray electrons. Since we are making measurements in the 1-MHz region, this corresponds to cosmic ray electrons with energies of a few hundred MeV.

Figure 2 shows a comparison of the results of calculations made from the radio measurements and the direct measurements of the flux of cosmic ray electrons. It is similar to Figure 1 of Dr. Cline's presentation. The circle located inside the shaded area represents the calculation that we

made from an analysis of the radio spectrum, and it is appreciably higher than the electron fluxes measured during balloon and satellite experiments. The dashed line indicates the slope of the electron spectrum that we would deduce in this energy range, and the shaded box is an indication of the precision of the measurement at this point. The uncertainty shown by the error box is not due primarily to the inaccuracy of the measurement but, at present, is due to our uncertainty of the magnitude of the galactic magnetic field and our estimate of the scale over which we are sampling the radiation.

To give you an idea of this problem, if we were to increase the estimate of the scale in the local spiral arm that we feel we are sampling and raise the galactic magnetic field from 3 to 5 μG, we would be moving the measured point down to the bottom of the error box. But, the important point to note is that the calculation from the radio data indicates that the true interstellar flux of low-energy cosmic ray electrons is appreciably higher than what one measures in the neighborhood of the Sun. Presumably, this is the result of the effects of residual solar modulation of the electrons that get into the solar system.

In summary, we were able to determine that the density of the thermal electrons in the interstellar gas is consistent with current theories and other techniques of measurement; whereas we found the measurement of the cosmic ray electron flux to be appreciably higher than what is measured at the Earth. The measurement from the radio data is presumably more representative of the true interstellar value outside the solar influence.

CHAIRMAN:

I think that this paper and Dr. Fainberg's paper show some of the really exciting results that are coming out of the RAE program. The spacecraft continues to operate well and will eventually be augmented by another one, in lunar orbit.

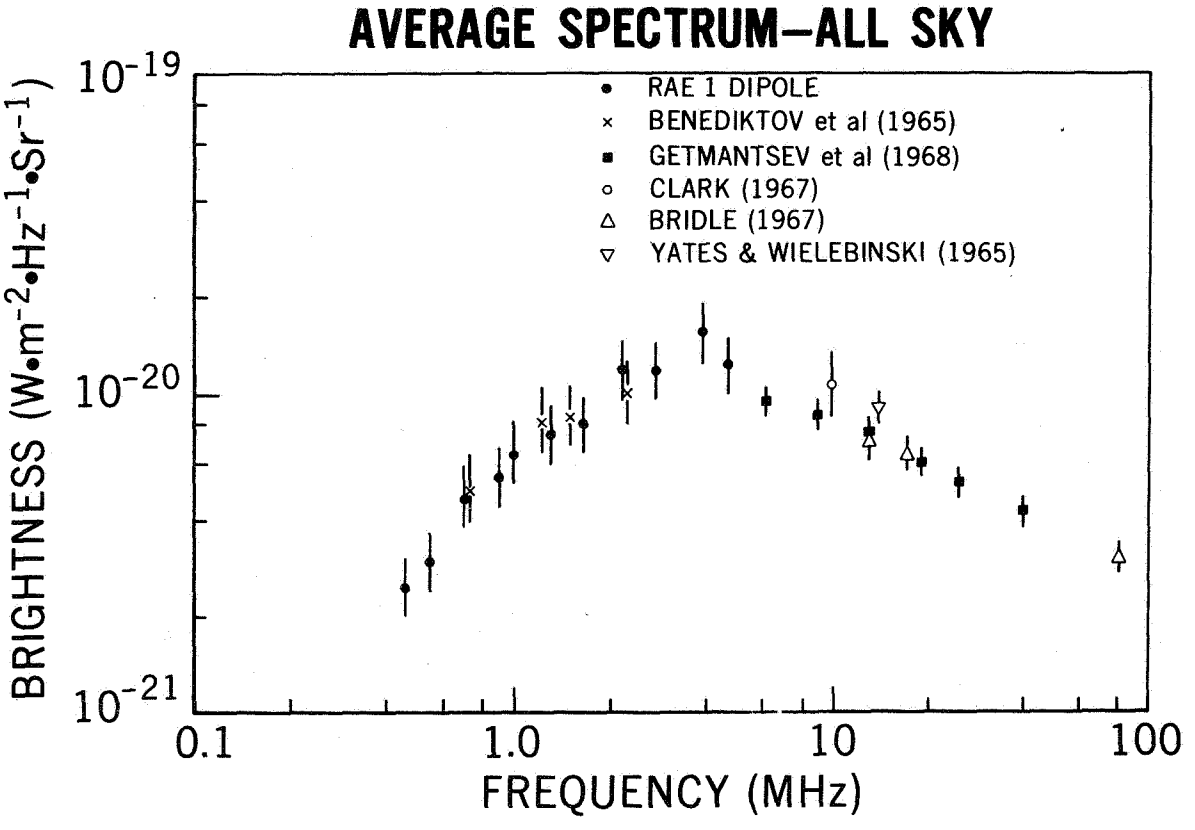

AVERAGE SPECTRUM—ALL SKY
RAE 1 DIPOLE
BENEDIKTOV et al (1965)
GETMANTSEV et al (1968)
CLARK (1967)
BRIDLE (1967)
YATES & WIELEBINSKI (1965)
BRIGHTNESS (W•m-2•Hz-1•Sr-1)
10-19
10-20
10-21
0.1
1.0
10
100
FREQUENCY (MHz)

Figure 1

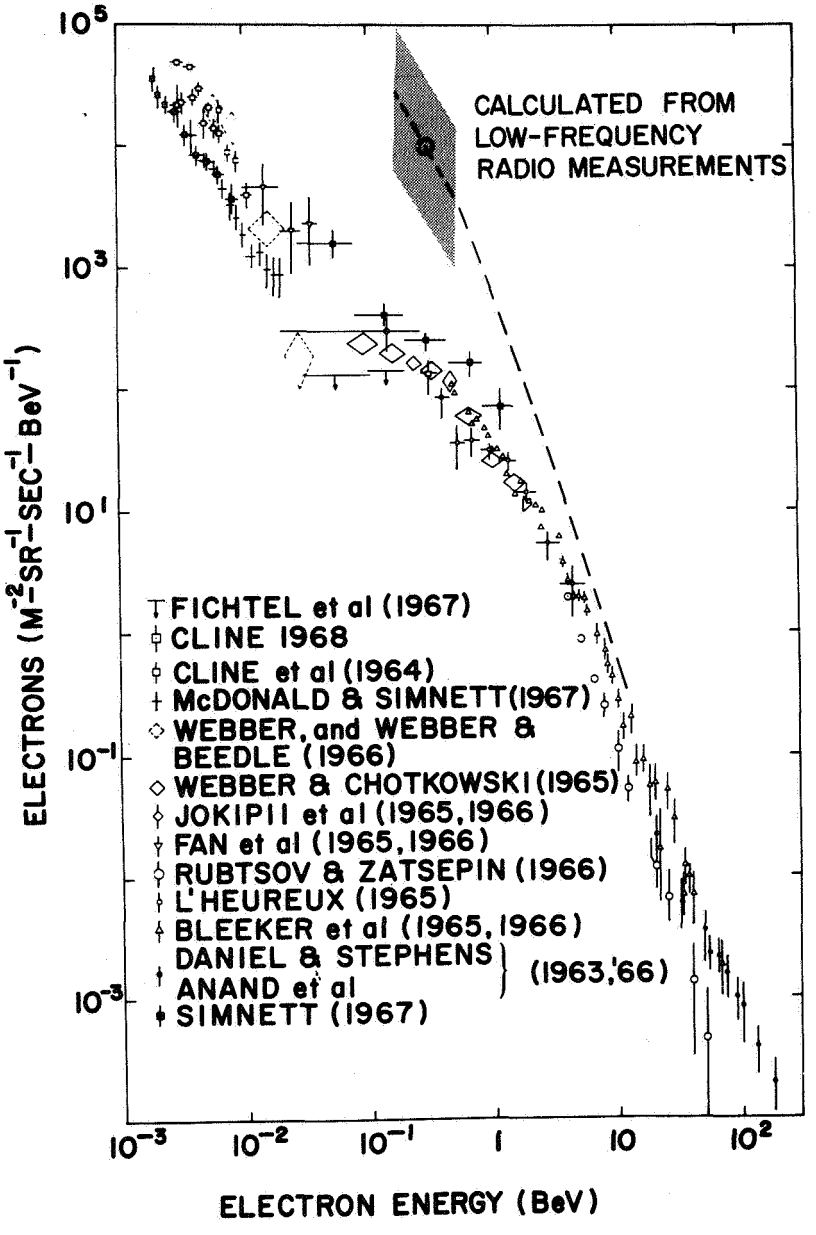

CALCULATED FROM
LOW-FREQUENCY
RADIO MEASUREMENTS
FICHTEL et al (1967)
CLINE 1968
CLINE et al (1964)
McDONALD & SIMNETT(1967)
WEBBER, and WEBBER & BEEDLE (1966)
WEBBER & CHOTKOWSKI (1965)
JOKIPII et al (1965, 1966)
FAN et al (1965, 1966)
RUBTSOV & ZATSEPIN (1966)
L'HEUREUX (1965)
BLEEKER et al (1965, 1966)
DANIEL & STEPHENS
ANAND et al
(1963,'66)
SIMNETT (1967)
ELECTRONS (M⁻²SR⁻¹SEC⁻¹BeV⁻¹)
ELECTRON ENERGY (BeV)
10⁵
10³
10¹
10⁻¹
10⁻³
10⁻³
10⁻²
10⁻¹
1
10
10²

Figure 2

LOW ENERGY COSMIC RAYS IN THE SOLAR SYSTEM

James H. Kinsey

CHAIRMAN:

We next hear from Dr. James Kinsey on low energy cosmic rays in the solar system.

DR. KINSEY:

Until about the last year, the cosmic ray spectrum at low energies had been very difficult to understand. Most observers have believed that there was a relative minimum in the low energy spectrum at around 30 MeV during times of very quiet solar activity.

In Figure 1, I give four representative 4-day spectra taken at different levels of solar activity. All of these data are from the two IMP 4 Goddard experiments: Dr. Hagge's low energy, solid-state detector telescope, and Dr. McDonald's tried and true double scintillator telescope. The first period for June is very near the launch of the satellite, just after a very large flare, and as one can see we have, for both protons and alphas, a monotonically decreasing spectrum. This is exactly what we would expect to see. At a slightly later date in June, the flare particles are leaving the solar system, and we end up with the minimum I spoke of, which is, in this case, at the higher energy end of the spectrum at about 80 MeV per nucleon. Finally, the last two frames, for 19 to 22 July and 5 to 9 September, show spectra for very quiet solar periods where the minimum is down as low as around 10 MeV.

The foregoing was very suggestive, and we have come to the conclusion that this indicates that the sun is actually producing a continuous or quasi-continuous flux of energetic particles, so that what we are viewing during quiet times is not the quiet, modulated spectrum of galactic cosmic rays, but is instead a mixture of the low energy tail of the modulated galactic spectrum and the solar spectrum at these low energies.

Figure 2 shows a model of what we think is happening. Over a long period of time, the period of a year for instance, one would not expect the modulated galactic spectrum to vary too much, whereas the solar particles, which are produced by flares in various active regions of the sun, would have a very large dynamic range of fluxes. As we know, flares can produce changes in particles fluxes at 20 MeV of as much as 6 orders of magnitude above the background.

We contend that what we are seeing is the "riding up," if you will, of the various levels of solar particle fluxes on the galactic spectrum, thus producing this minimum which heretofore had been thought to be caused by some peculiar property of the interplanetary medium. Figure 2 shows that we really should be able to make a couple of predictions about what we might expect to see and how we might unravel this phenomenon.

What we predict is that there should be, first of all, no correlation between the low energy fluxes and the high energy fluxes when we have the situation where the minimum is visible in the spectral region of interest. Secondly, we expect that the position in energy of the minimum would be very closely correlated with the level of the low energy flux. On the other hand, we expect that the high energy flux level would show essentially no correlation with the position of this flux minimum. Finally, we expect that there would be a positive correlation between the flux at minimum energy and the minimum energy itself.

In Figure 3, we show the correlations for the differential energy flux for protons. Let me add that the results for alpha particles are almost identical. The differential flux at around 76 MeV plotted against the differential flux at around 5 MeV, down at the low energy end of our spectrum, is shown in Figure 3(a). As you will note, during very low solar activity, there is no correlation. We find that at about 10^{-1} in our units (number of particles per second per square centimeter per steradian per MeV), there begins to be a correlation. Now, this is where the solar particles are beginning to dominate. So the first thing that we predicted certainly seems to be true in this region: There is no correlation between the low and the high energies when the relative minimum is present. The second prediction is that we should have very good correlation between the low energy flux and the energy and a minimum flux. In Figure 3(b), we have the 5-MeV differential flux plotted against the energy at minimum flux, and we see a very good correlation. The third prediction was that we would have no correlation between the high energy flux in our energy interval and the energy at minimum flux, and I think this is shown very clearly in Figure 3(c). Finally, in Figure 3(d), the rather trivially obvious thing is that the minimum flux itself should have a positive correlation with the energy at minimum flux and indeed, this is what we do see. One can go on, which I am not going to do here, and show that there is good correlation between the data presented here and other forms of solar activity.

DR. KRIMIGIS:

Have you considered our recent findings of a negative radial gradient of cosmic rays at low energies? Your results are certainly consistent with this.

DR. KINSEY:

In the present study, we have not considered the negative gradient. I agree, however, that the two findings are consistent.

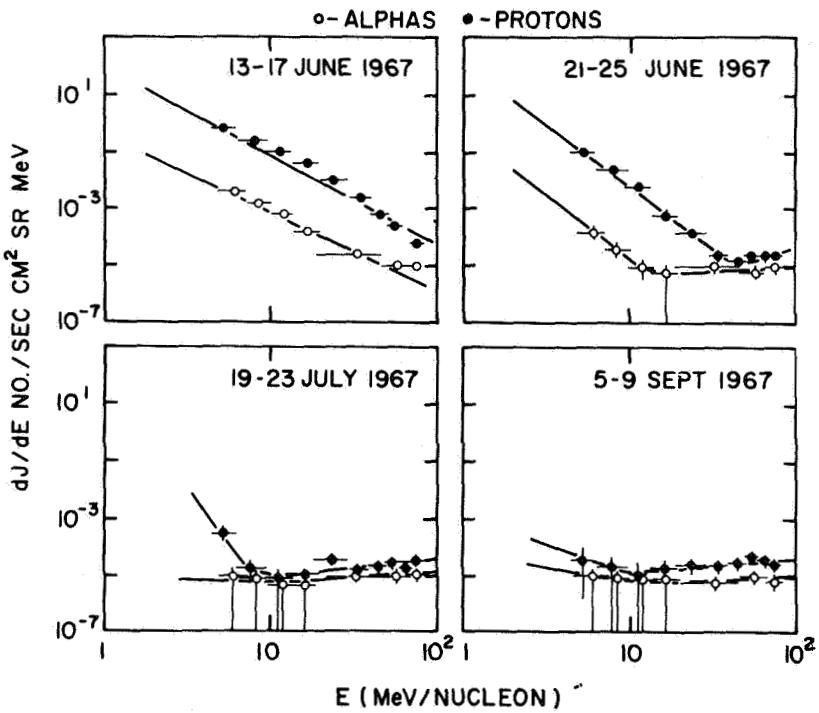


Figure 1 – IMP 4 particle spectra.

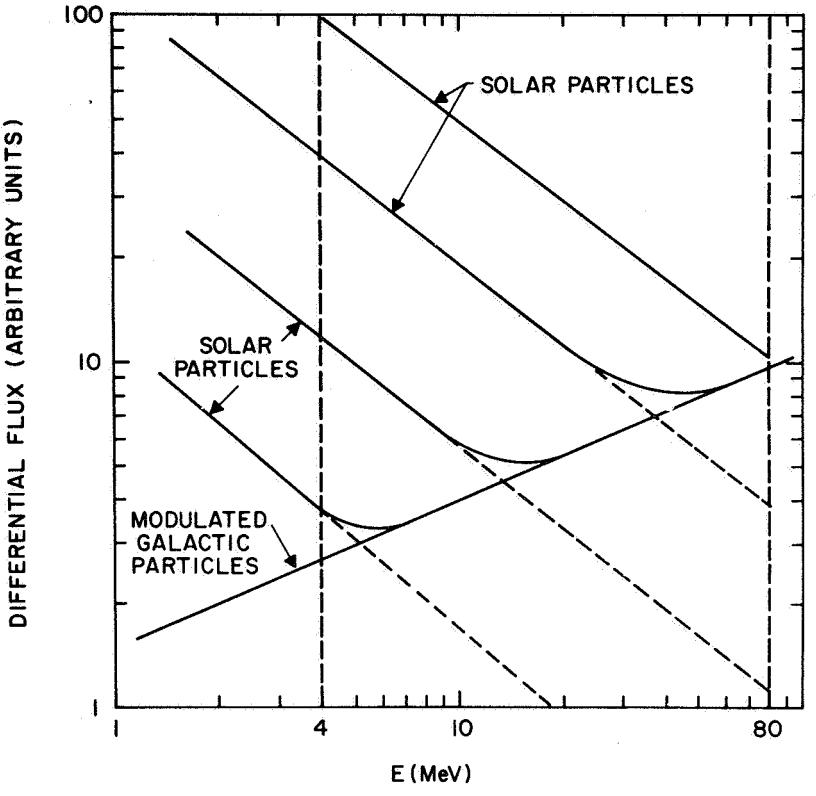

100
10
1
DIFFERENTIAL FLUX (ARBITRARY UNITS)
SOLAR PARTICLES
SOLAR PARTICLES
MODULATED GALACTIC PARTICLES
1
4
10
80
E (MeV)

Figure 2

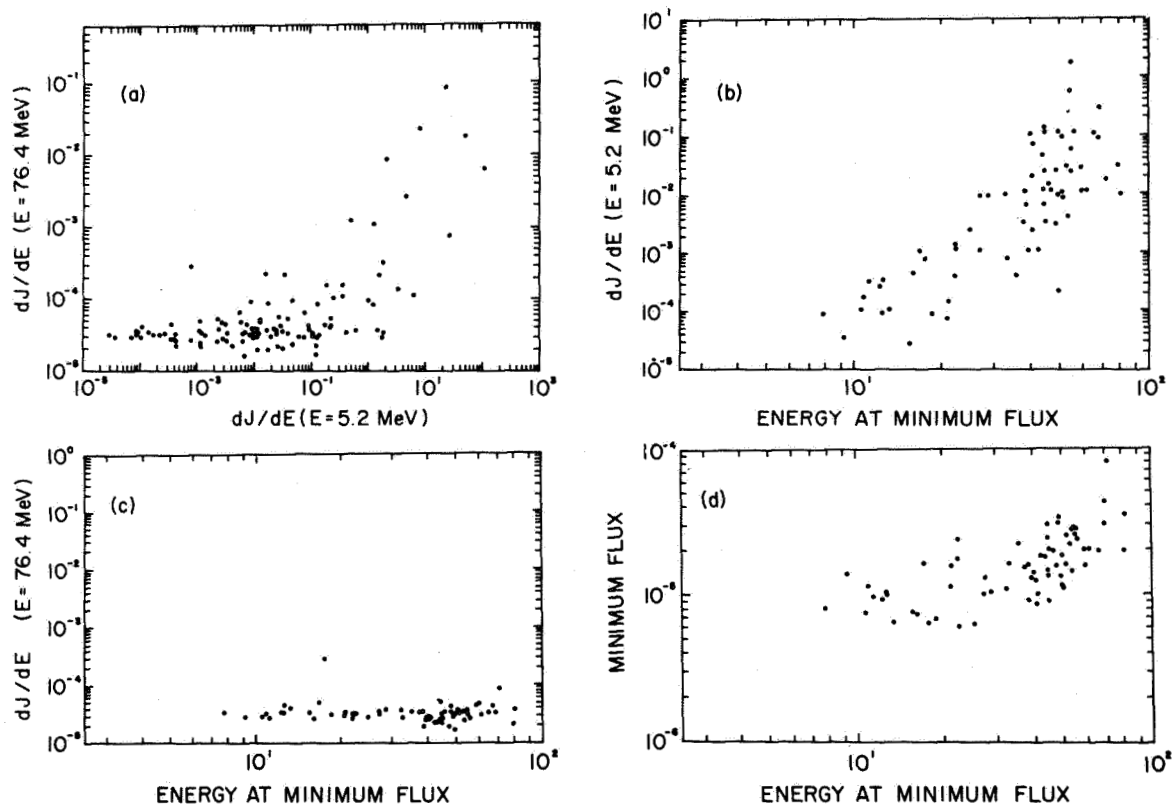

(a)
dJ/dE (E = 76.4 MeV)
dJ/dE(E=5.2 MeV)
(b)
dJ/dE (E = 5.2 MeV)
ENERGY AT MINIMUM FLUX
(c)
dJ/dE (E = 76.4 MeV)
ENERGY AT MINIMUM FLUX
(d)
MINIMUM FLUX
ENERGY AT MINIMUM FLUX

Figure 3

PULSARS AND THE ORIGIN OF COSMIC RAYS

Reuven Ramaty

CHAIRMAN:

The next paper concerns pulsars and the origin of cosmic rays. It will be given by Dr. Reuven Ramaty.

DR. RAMATY:

Studies of the origin of cosmic rays indicate that, except for very low energies below about 40 MeV per nucleon and the very high energies above about 10^{11} MeV, these particles are all of galactic origin. The most energetic events in the galaxy seem to be supernova explosions and possibly the galactic center.

The identification of the explosive events that are the most important contributors to the observed cosmic ray flux at the earth would be of great interest, both for the understanding of the origin of the cosmic rays and for the evaluation of some of their principal properties, such as age, matter-traversal, and degree of anisotropy.

Even if the positions and ages of the cosmic ray sources were known, cosmic ray propagation in the galaxy is a complicated problem. As a simplifying condition, we have assumed that the cosmic rays diffuse isotropically, the diffusion mean free pass, λ, being a free parameter which is in principle determined by the interactions of the cosmic rays with magnetic irregularities in interstellar space.

Considering the cosmic ray energy density at the earth from a single source, in Figure 1 we show the energy density at the earth, as a function of time, from a source of total energy in relativistic particles of 10^{51} ergs and situated at a distance, R, of 100 pc from the earth. The time t_m is the time to maximum for diffusion and is given by $R^2/2\lambda c$, where c is the velocity of light.

As one can see, for a single source of particles, the energy density as a function of time rises rapidly to a maximum and then decays approximately like $t^{-3/2}$.

In the past, except for a few historic supernovae and the observations of radio shells of supernova remnants, the distances and ages of most of the cosmic ray sources were not known, and hence average distributions were used. With the discovery of pulsars, however, a more realistic picture of the positions and times of the supernova explosions can be obtained.

Assuming that pulsars are rotating neutron stars that slowed down while dissipating their rotational energy, we can estimate their ages (from the measured periods and rate of change of the period) and distances (from the observed dispersion of the radio pulses). The basic assumption is that the pulsars are the end products of the supernova explosions and hence the distances and ages of the cosmic ray sources are the same as those of the pulsars.

These are summarized in Figure 2. The dots denote the totality of the observed pulsars; the distances are given in parsecs and the ages in years.

Also shown in this figure are a number of historical observations of supernova explosions. Among these, only the Crab Nebula is an observed pulsar.

The contour lines indicate the location in space-time of a source which would contribute at the earth the indicated amount of energy density. Since the observed energy density of cosmic rays at the earth is on the order of 10^{-12} ergs cm^{-3}, from Figure 2 we can make the statement that we are confronted by a fairly complete set of cosmic sources, which when given a total energy of a few times 10^{51} ergs per event can produce the observed cosmic ray energy density at the earth.

With a mean free path ranging from a few hundredths of a parsec to maybe a parsec, the properties of the cosmic rays, their age, matter-traversal, and degree of anisotropy are all consistent with observations and with independent estimates.

Finally, there exists the distinct possibility that for very short path lengths a few of the close-by sources, such as CP0950, shown at a distance of 60 pc in Figure 2, could contribute the bulk of the observed cosmic rays at the earth. This then would lead to a situation in which the cosmic rays observed at the earth are not representative of the distribution of energetic particles in the galaxy, but represent only the particle population in a limited portion of space surrounding the solar system.

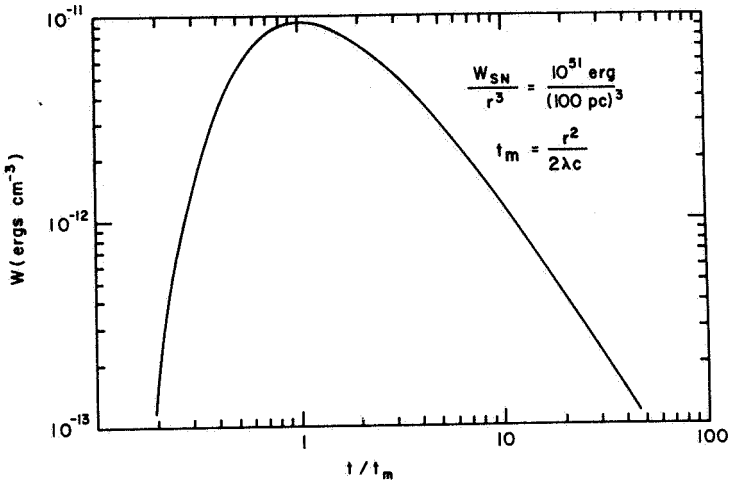


Figure 1

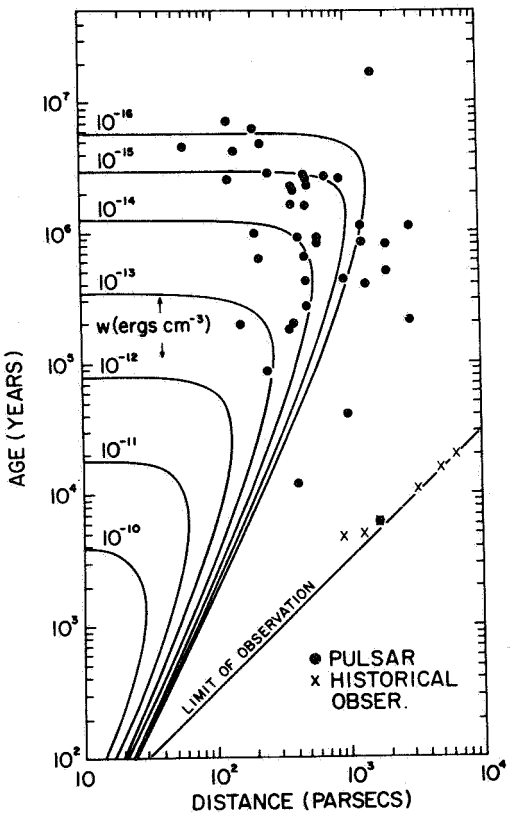


Figure 2

ULTRA-HIGH-ENERGY COSMIC RAYS

Floyd W. Stecker

CHAIRMAN:

The next speaker is Dr. Floyd Stecker who will discuss ultra-high-energy cosmic rays.

DR. STECKER:

Cosmic rays are now known to exist with energies up to 3×10^{20} eV. This maximum energy is $\sim 10^{10}$ times higher than the highest energy particles man has yet obtained with present accelerators. It represents enough energy, in a single cosmic ray particle the size of an atomic nucleus, to lift a person of my weight to a height of more than 2 inches off the ground.

Now the question naturally arises: Where do these cosmic ray particles, which are obviously the result of tremendously powerful cosmic forces, originate?

Figure 1 shows the total energy spectrum of cosmic rays as presently known. We can get a hint from this figure by noting that I have divided up the total spectrum into two components: a galactic component that is the component Dr. Ramaty was discussing and, above about 10^{18} eV, an extragalactic component with a flatter energy spectrum, as shown by the recent data of the Sydney air shower group.

There are three reasons for believing that the ultra-high-energy cosmic rays (those above $\sim 10^{18}$ eV) are of extragalactic origin:

(1) The energy spectrum indicates an excess of these cosmic rays above what one would expect by extrapolating from the lower energy part of the total spectrum.

(2) These cosmic rays have such high energies that they cannot be contained within the magnetic field of the galaxy.

(3) Possibly the most important reason, these cosmic rays are arriving at the Earth from random directions over the whole celestial sphere. Thus, their arrival directions are not concentrated within the plane of the Milky Way.

My question is, from what distances are these cosmic rays, which are assumed to be of extragalactic origin, reaching us, i.e., how far away are their sources?

To answer this question, I have made use of one of the most important recent discoveries in astrophysics, namely the discovery that the whole universe seems to be bathed in a field of thermal radiation left over as a remnant from the original "big bang" of the universe and that this radiation field contains thermal photons at a density of approximately 10^3 cm^{-3}, having an average energy of about 10^{-3} eV/photon.

Now, we know from the laws of special relativity that in the rest frame, i.e., the inertial frame of one of these ultra-high-energy cosmic rays, the thermal photons appear to be high-energy γ rays with enough energy to interact with the cosmic ray and convert some of its energy into matter in the form of pions. By calculating the lifetime of these cosmic rays against this energy loss attenuation process, we can thus determine an upper limit to the distance of the source of these cosmic rays. Since these cosmic rays are traveling at very nearly the speed of light, the maximum distance in light-years at which they could have originated is equal to the lifetime of the cosmic rays in years.

Figure 2 shows the results of the lifetime calculation for the cosmic ray nucleus in years versus total energy. If we look at the highest known energy, about 3×10^{20} eV, we see that cosmic ray protons of this energy (H nuclei) can exist for only 10^8 years before being attenuated by this process. Hence, they must be reaching us from sources 10^8 light-years distant or closer, a distance that is much smaller than the visible radius of the universe (10^{10} light-years, indicated by the horizontal dashed line). Therefore, the sources of these cosmic rays are relatively local.

I have also considered that the remnant thermal photons can interact with high-energy cosmic ray nuclei, breaking them up into individual nucleon constituents, and I have calculated the lifetime for the breakup of nuclei due to this attenuation process. At an energy of 3×10^{20} eV, light cosmic ray nuclei such as helium can survive for only about 2×10^6 years before being destroyed. However, heavy cosmic ray nuclei of the same total energy can exist for the same order of lifetime as the protons (10^8 years).

Thus, we can reach the following conclusions: It appears that ultra-high-energy cosmic rays are of extragalactic origin; that they originate in relatively local extragalactic sources (10^8 light-years or less); and that the constituent nuclei of these cosmic rays must be either protons or heavy nuclei such as iron.

DR. CLARK:

I was curious about the higher energy cosmic rays, those having about 10^{22} eV. In Figure 2, it looked like you had a minimum lifetime near there, so presumably cosmic rays of higher energies would have longer lifetimes, and cosmic rays having energies around 10^{20} or 10^{21} eV might even essentially be secondaries of high-energy primaries.

DR. STECKER:

For the protons (H nuclei), the lifetime curve levels off to an asymptotic value which is close to 10^8 years at energies that are higher than the range shown in Figure 2. So, this actually puts an upper limit on the distance of all cosmic rays above 10^{21} eV. I have not depicted this in the figure because there are not any cosmic rays known above 3×10^{20} eV at present.

DR. CLARK:

What about the α's and the heavier particles?

DR. STECKER:

For the α's (He nuclei) and the heavier particles (Fe), the lifetimes increase at energies higher than those shown here. The minimum in the curves is due to a resonance in the photodisintegration cross section which occurs at a few MeV when measured in the rest systems of the particles.

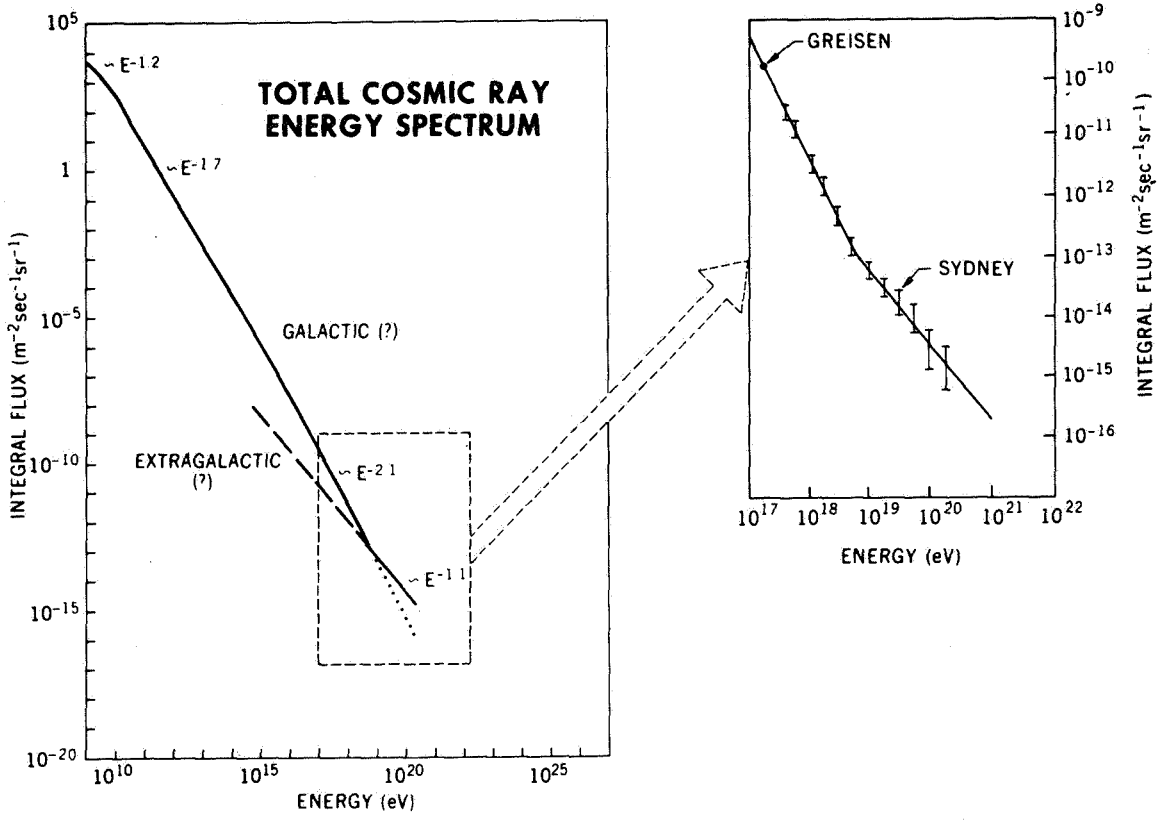

TOTAL COSMIC RAY ENERGY SPECTRUM
~E^-1.2
~E^-1.7
GALACTIC (?)
EXTRAGALACTIC (?)
~E^-2.1
~E^-1.1
INTEGRAL FLUX (m^-2sec^-1sr^-1)
ENERGY (eV)
GREISEN
SYDNEY
INTEGRAL FLUX (m^-2sec^-1sr^-1)
ENERGY (eV)

Figure 1

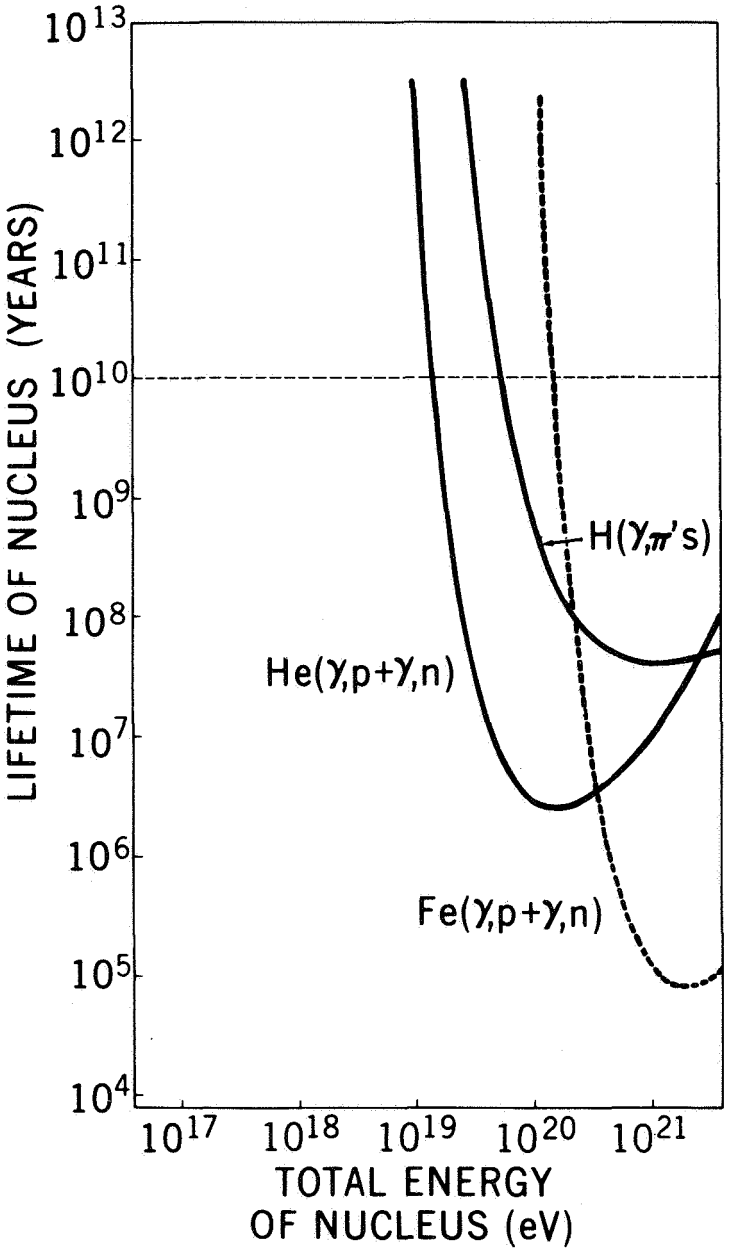

LIFETIME OF NUCLEUS (YEARS)
10^{13}
10^{12}
10^{11}
10^{10}
10^{9}
10^{8}
10^{7}
10^{6}
10^{5}
10^{4}
H(γ,π's)
He(γ,p+γ,n)
Fe(γ,p+γ,n)
10^{17}
10^{18}
10^{19}
10^{20}
10^{21}
TOTAL ENERGY
OF NUCLEUS (eV)

Figure 2

MOLECULES IN INTERSTELLAR SPACE

Louis J. Stief

CHAIRMAN:

We turn now to the subject of molecules in interstellar space, and we will hear from Dr. Louis Stief.

DR. STIEF:

One of the most rapidly advancing areas in space science has been the detection of polyatomic molecules in the interstellar medium by means of radio astronomy. To give you a personal measure of the pace of these new observations, all the molecules so far observed–and there have been only three: ammonia, water, and formaldehyde–have been observed since I came to Goddard just a year ago.

Now the detection of formaldehyde especially was somewhat surprising because it is a reasonably complex and reactive molecule. The question immediately arises: What is the lifetime of this molecule in the interstellar radiation field. How does it compare with the other molecules we have observed?

To answer such questions we need to know three things. First, we need to know the photon flux in interstellar space, and this is essentially an observational problem. Second, we need to know the absorption cross section for the molecules in question, and this is a laboratory study in spectroscopy. Finally, we need to know the efficiency of the photodecomposition, the fraction of molecules that decomposes upon photon absorption. This is a laboratory problem in photochemistry in which we need to know not only how many molecules absorb photons but what the consequence of that absorption is.

Concerning the first question, it is pretty certain now that the flux of interstellar radiation is fairly constant, within a factor of 2, throughout interstellar space and throughout the spectral region from 2400 Å down to 912 Å, with a few important exceptions such as cloud areas.

Now concerning the other two questions, the absorption spectra and the photochemical decomposition of both ammonia and water are very well investigated and understood, even down to the short wavelengths. However, this is not the case for formaldehyde.

Not a single study of the photodecomposition of formaldehyde has been reported below 2000 Å, and even the available absorption spectra are quite poor, quite qualitative, and very limited in value.

Dr. Mentall and Mr. Gentieu in the Astrochemistry Section are now in the process of making quantitative photoelectric measurements of the formaldehyde absorption, and their initial results are shown in Figure 1. This covers the region from 900 Å up to about 1600 Å. The thing to note here is the underlying weak continuum which was not evident in the older, mainly photographic, data and which, as we will see in a moment, has very important implications for the lifetime calculation.

The other point worth noting is the strength of the absorptions. In the units I am using here, reciprocal centimeter atmospheres, they run from 200 to 2000. Now the absorptions in the near ultraviolet, where all the work has been done, never reach more than 2. So that with these absorptions being up by a factor of 100 or 1000, it is evident that absorption below 1600 Å is going to determine entirely the photodecomposition and the lifetime of the molecule. Yet, this is just where we are completely lacking in information.

For this reason Dr. Glicker and I undertook a laboratory study of the photodecomposition of formaldehyde vapor and we used the 1236-Å krypton resonance line and the 1470-Å xenon resonance line. These are readily generated in the laboratory, and they allow us to look at the effect when these wavelengths of light are absorbed by the formaldehyde molecule. Our summary of results from a large number of experiments is:

$$HCHO + h\nu \xrightarrow{\phi=0.5} H_2 + CO$$

$$HCHO + h\nu \xrightarrow{\phi=0.5} 2H + CO$$

$$HCHO + h\nu \not\longrightarrow H + HCO$$

These results did not depend on the wavelength or intensity of the light or on other parameters such as pressure. We used deuterated and undeuterated formaldehyde to distinguish the first two processes where hydrogen is formed. In a direct molecular process, this would not lead to mixing of hydrogen isotopes. The second case, the formation of hydrogen atoms which subsequently end up in the laboratory as H_2, would lead to a scrambling of the isotopes.

Now, two results here are important for our purposes. One is that the sum of the efficiencies of the first two processes is unity; that is, all the molecules which absorb do indeed decompose. We need to know this to make a good lifetime calculation.

The second point is that the third process, which is one of the dominant processes above 2000 Å, is unimportant here. Therefore, the HCO radical would not be expected to be present in interstellar medium as a result of photodecomposition of formaldehyde.

Based on the photon flux, the absorption cross section, and the knowledge of the efficiency of the decompositions, Dr. Donn and I have calculated the lifetime of these interstellar molecules, and these are shown in Table 1.

Table 1

Lifetime of Interstellar Molecules

Molecule	Fragments	Spectrum	τ (years)
H_2O	$H + OH$ $H_2 + O$	Continuum only $\lambda < 1850$ Å	10^3
NH_3	$H + NH_2$ $H_2 + NH$	Continuum with bands $\lambda < 2200$ Å	10^3
HCHO	$2H + CO$ $H_2 + CO$	Continuum with bands $\lambda < 1600$ Å	10^3

Use of the older photographic data for formaldehyde would have led to a value an order of magnitude or two higher than that given here. This is one of the important aspects of having an integrated program where we can look at the laboratory as well as the observational problems. However, if we use what we now believe to be good quantitative data, the lifetimes for all molecules are of the order of 1000 years.

You see, there is some similarity in the way the molecules fragment. They all have unit efficiency for decomposition. Water absorption is mainly continuous, whereas the last two, ammonia and formaldehyde, are continuous with bands.

Now these lifetimes, of course, will be longer in clouds where there is obscuration of the interstellar radiation. They could be as high as 10^4 or

10^5 years. And this, of course, is the region where these molecules have been mainly observed by radio astronomy. However, these are still rather short lifetimes on the galactic time scale. So, what are required are efficient mechanisms of molecule formation to explain the abundances of these complex molecules in space.

MEMBER OF THE AUDIENCE:

Are there other molecules which you would expect to find similarly?

DR. STIEF:

There are other molecules that have been looked for and not found, molecules of comparable complexity that contain the elements carbon, hydrogen, and oxygen.

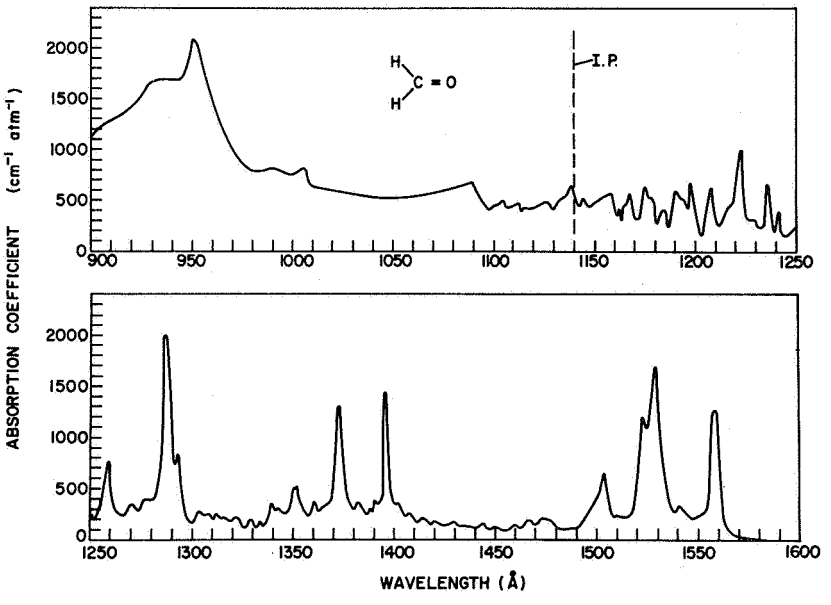


Figure 1

INTERSTELLAR EXTINCTION

Theodore Stecher

CHAIRMAN:

Still on the subject of molecules and atoms in space is Mr. Theodore Stecher's paper on interstellar extinction.

MR. STECHER:

Interstellar extinction is the attenuation of light, as it passes through interstellar media, by absorption or scattering. It is sometimes called reddening because it is inversely proportional to wavelength; that is, a distant star behind a cloud will appear redder than an identical but closer one. One uses this fact to determine the extinction.

Figure 1 shows a curve which I published four years ago. The vertical axis is extinction in magnitudes, which is the astronomer's logarithmic scale, and the horizontal axis is $1/\lambda$, with the wavelength in microns. The ground observations are up to $3\mu^{-1}$. The rest we got with a rocket which just went up and spun around; we discovered what stars we had looked at after the fact. It is a hard way to do astronomy.

I want to call your attention to what looks like a bump in the center and to the general shape of the curve at shorter wavelengths.

Figure 2 shows the same points with two theoretical curves on it. The solid curve, which is due to van der Hulst, is the extinction you would get from ice particles of the order of the wavelength of light. The broken curve would be due to graphite particles, the hump being produced by transition of the π-electron to the conduction band. The measurements of the complex dielectric constant for graphite came out at just about the same time as our rocket measurements.

At shorter wavelengths, atomic and molecular gases can produce the absorption that is necessary, but this is not true in the visible region. In the Thirties, we showed quite clearly that it was necessary to have a Mie-type particle, a particle the same size as the wavelength of light that was being extinguished. Above about $11\mu^{-1}$, where atomic gases completely dominate, atomic hydrogen is the chief cause of extinction.

My colleagues and I decided to believe the observations. We constructed a model of an interstellar grain, a piece of graphite sealed up a

million times larger than the real interstellar particle. We can understand it, make measurements on it, and perform calculations. We wanted to see how far we could push it. My colleagues, incidentally, were Donn, Wickramasinghe, Hudson, Krishnaswami, and David Williams. We put together what we think is a fairly plausible story.

In late type stars, where the partial pressure of carbon exceeds the vapor pressure of graphite and there are a few ions around, we can get nucleation, that is, enough particles coming together to form a little nucleus which will then grow by crystal growth until it gets large enough for radiation pressure to act upon it. When it gets that large, it starts to be blown out of the star, taking some gas with it. This gives us a method of mass loss for these late type stars which is observed and which hydrodynamic theories have failed to produce.

The particle can be aligned by diamagnetic relaxation in the stellar field outside the star, and the cloud there will explain the infrared excesses which are observed in these stars. It will produce the intrinsic polarization that Mira stars show, and in the general interstellar media, it will then again produce interstellar polarization and extinction.

It is a very strongly anisotropic crystal. That is, the conductivity varies by a factor of 10^3 to 10^5 in different directions; so, projection of an electric vector in two directions will produce polarization. We went on and investigated a number of other properties of it, too, and did some speculation on X-ray scattering, a theory of chemical exchange processes, and the formation of diatomic molecules–we had not discovered polyatomic ones at that time.

We put together a story which is quite plausible, but, as in all astrophysical matters, we never had quite enough information, and we did not have enough observations to really feel completely happy. So, in this case, what one would like to do is to go back and show that the bump is really there in the extinction curve; one could, after all, draw a straight line through the points and wash out the bump.

Last year I had the opportunity to observe stars of my choosing with a pointing rocket employing the strap system, and the results are shown in Figure 3. The stars were Epsilon and Zeta Persei. Zeta Persei is obscured–behind the cloud; Epsilon is in front of the cloud. Both stars are of similar spectral type. The extinction versus inverse wavelength plot clearly shows my bump.

Concluding, it appears that the stars are smoking like candles, and the confirmation of my reddening curve has me tickled pink.

CHAIRMAN:

Is there any question or comment on Stecher's bump?

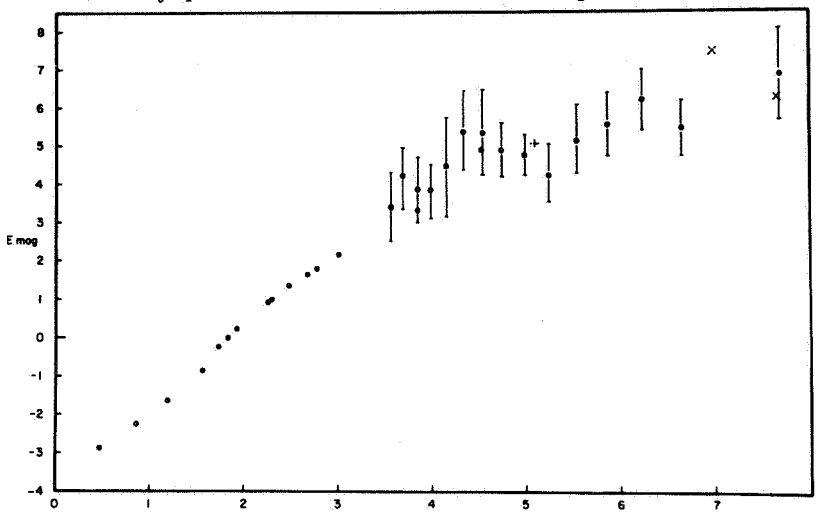


Figure 1

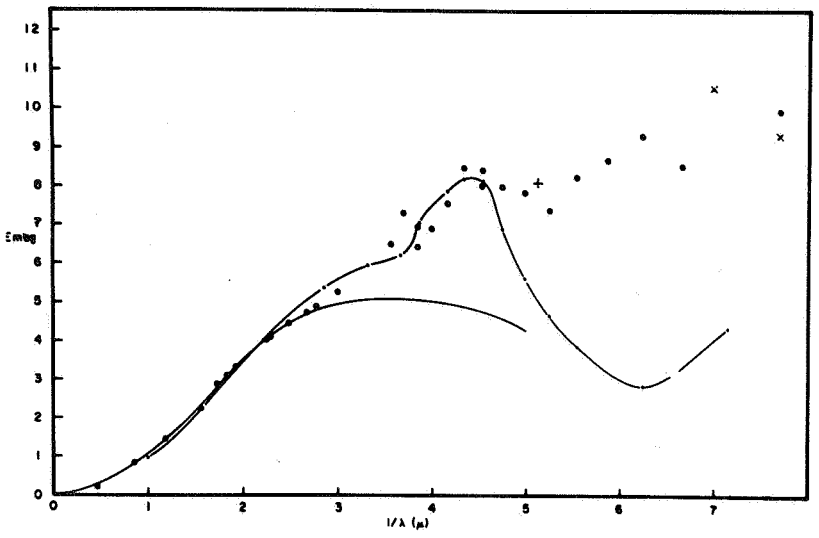


Figure 2

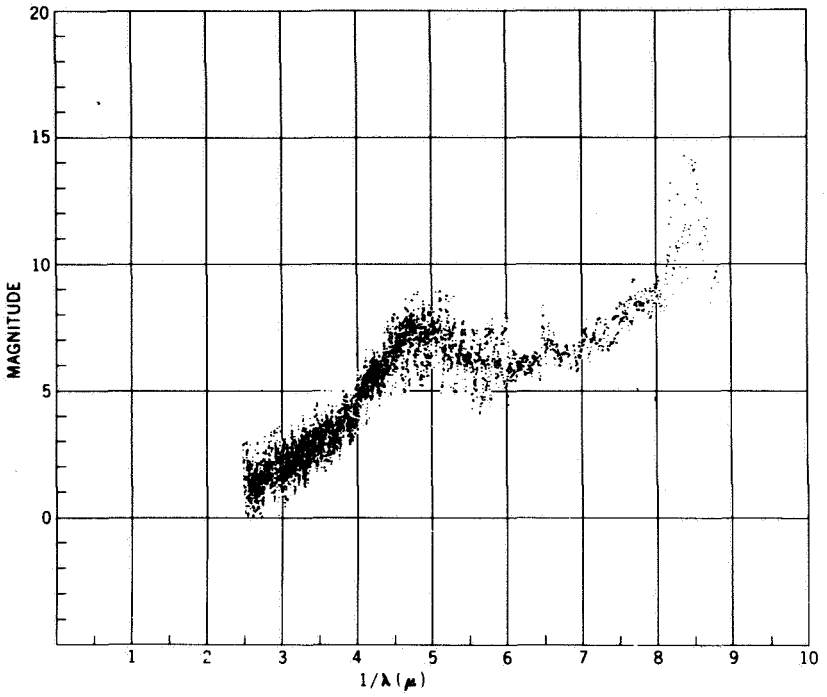


Figure 3

OXYGEN PRESSURE DURING THE FORMATION OF METEORITES

Louis S. Walter

CHAIRMAN:

Next we turn to a question concerning the formation of meteorites. Oxygen pressure is a matter of considerable importance during this formation, and this will be described to us by Dr. Louis Walter.

DR. WALTER:

Meteorites have been dated at about 4.5 billion years. They are the source of just about all the information we have on the formational processes in the beginning of the solar system, the beginning of the planetary system, and the planets. One of the important parameters that we should know in determining the "condensation processes" in the formation of the planets is the oxygen pressure. Until recently the data that we had have come from the equilibrium experiments made at temperatures around 1200° C. It has been impossible to get data down around 800° C. The reason the 800° C temperature is important is the following.

The stony meteorites, which most of our theories are concerned with, are equilibrated. That means that the chemical composition of the various minerals from one part to the other is the same. And the temperature of equilibration has been determined by Van Shmus and Kaufmann to be approximately 800° C, meaning that the meteorites were exposed to these temperatures at about 4.5 billion years.

Now if we are going to derive values for oxygen pressure, we are going to have to have pertinent data at 800° C. Since the data that are pertinent have only been those at around 1200° C, we have been in sort of a fix.

I decided to determine the oxygen pressure by a rather novel technique in which high temperature electrolytes are used. I will show you, in Figure 1, that this electrolyte, zirconium oxide mixed with calcium oxide, is transparent to oxygen ions at temperatures above 600° C. On one side we used a buffer of known oxygen pressure, and we simply put the sample on the other side (Figure 1 is highly diagrammatic, as you probably realize) and measure the emf between the sample and the nickel-nickel oxide. As you can see, this is probably very similar to dry cells that you saw pictured in your physical chemistry courses (which may have been some time ago).

The whole thing is placed in the furnace with temperatures varying between 700° and 900° C, and the oxygen pressure, P_{O_2}, is determined by the equation given on the bottom of Figure 1.

Now this is a relative P_{O_2}; that is, it is the differential of oxygen pressure from the nickel-nickel oxide buffer, but fortunately, of course, we know what the oxygen pressure is for this as a function of temperature. What we are going to be interested in is the oxygen pressure as a function of temperature for this sample.

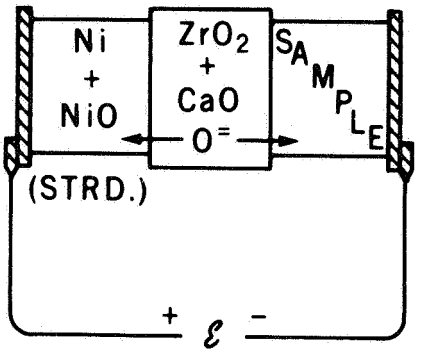
The sample is a meteorite. We take the meteorite *per se* and stick it in the furnace. So we are determining the oxygen pressure as a function of temperature for the meteorite itself.

There are four main types of equilibrated stony meteorites. Each one manifests its own type of oxygen pressure, so we are interested in the oxygen pressure of the four types.

Data for an enstatite-type chondrite are shown in Figure 2. I have plotted the emf in millivolts versus temperature in the upper graph, which also shows the data points obtained for this one run; then by the equation at the bottom of Figure 1, the log of P_{O_2} is plotted versus temperature on the abscissa in the lower graph. Now the data apply to an enstatite chondrite, as I said. There are three other types.

In Figure 3 you see the data, without the data points, for the four different meteorite types, LL, L, H, and E, in the temperature range determined, from 700° C to 900° C. So the data actually exist in this range. The data are particularly pertinent at 800° C, where we see that, for one group of meteorites, we have about 10^{-20}; for another, approximately 10^{-22}; and for the enstatite, the E-group meteorites, we have an oxygen pressure of about 10^{-24}.

These are valuable input parameters for the condensation theories, first of all; and second of all, they give us hydrogen-to-oxygen ratios which are about 50 times higher than those previously assumed for the primary solar nebula.



$$\Delta \text{LOG } Po_2 = - \frac{20.331\,\mathcal{E}}{T}$$

Figure 1

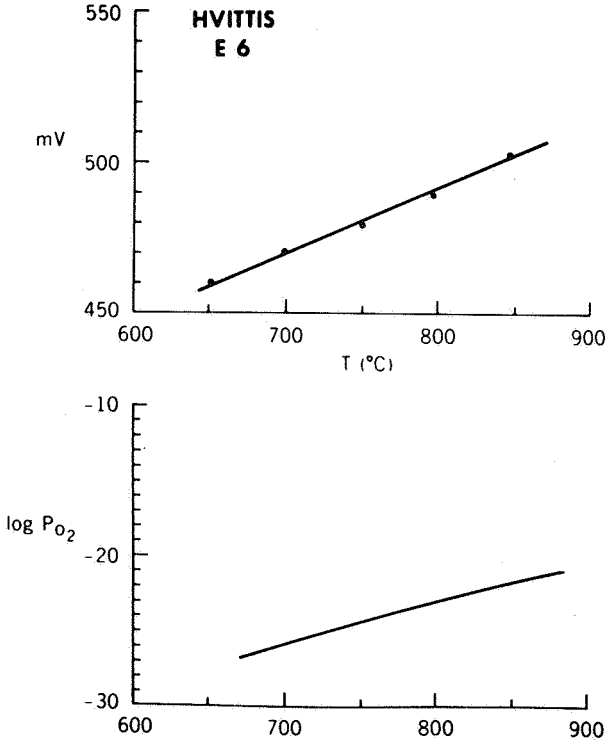


Figure 2

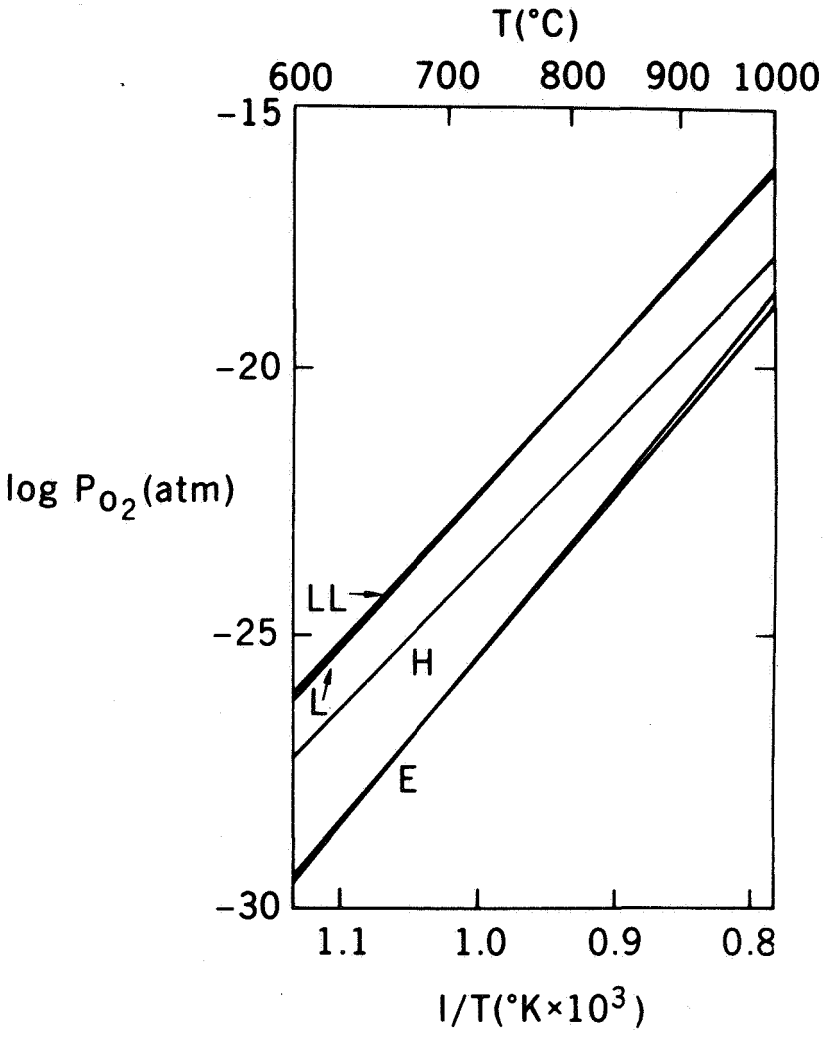

T(°C)
600
700
800
900
1000
-15
-20
-25
-30
log P_{O_2}(atm)
LL
H
L'
E
1.1
1.0
0.9
0.8
1/T(°K×10³)

Figure 3

PULSATIONS OF BLUE VARIABLE STARS

Richard B. Stothers

CHAIRMAN:

Next, Dr. Richard Stothers will discuss pulsations of blue variable stars.

DR. STOTHERS:

The idea that the release of subatomic energy inside stars might destabilize them and cause them to pulsate goes back to Eddington's work in 1918. Since that time, a few attempts have been made through theoretical investigations to find relevant stellar models, but, on the whole, these attempts have been rather unsuccessful in explaining observed variable stars.

It is well known that, since the sources of nuclear energy reside near the center of a star, a large amplitude of pulsation must develop in the stellar core if the star is to provide a sufficient input of energy during each cycle to overcome the heat leakage near the surface. Unfortunately, large core amplitudes can be induced only under conditions of high radiation pressure and low central condensation of the star. Such restrictive conditions immediately rule out the classical Delta Cephei variables as being energized by nuclear reactions because these stars have very large radii and therefore high central condensation. We are forced to look for more likely candidates among the bright blue stars of small radius and high mass, since high-mass stars have very high luminosities and hence high radiation pressures.

Figure 1 shows the results of a number of calculations for pulsationally unstable stars energized by nuclear reactions. Each point represents an unstable stellar model. Note that the objects with high central condensation ($\rho_c/\bar{\rho}$) require a high radiation pressure in order to be destabilized. Objects that fall below the approximate line delineated by the points, and which are not shown in the figure, are pulsationally stable. (The different symbols represent three basically different kinds of stellar models which will be applied to the explanation of three groups of observed variable stars: the Beta Cephei stars and two subgroups of Wolf-Rayet stars.)

Basically, two mechanisms can induce the required conditions for pulsational instability. The first mechanism is simply high mass and an approximately homogeneous chemical composition throughout the stellar interior.

Figure 2 shows a Hertzsprung-Russell diagram, which is a plot of stellar luminosity (increasing upward) against stellar surface temperature (increasing to the left). Two loci, along which the stellar models have a homogeneous chemical composition, are shown: hydrogen-burning stars on the right and helium-burning stars on the left. Each point again represents an unstable stellar model.

For example, one might hypothesize that the most massive objects near the hydrogen-burning main sequence (circles) may account for the high-luminosity group of Wolf-Rayet stars, a class of blue stars that manifests considerable instability and evidence of mass loss. It may well be that this mass loss is induced by the pulsational instability. Typical masses for the onset of this instability are calculated to be about 100 $M_{\odot}$, so that this group of luminous Wolf-Rayet stars, which appears to lie near the top of the hydrogen main sequence, would be quite massive.

The second way to induce pulsational instability is by reducing the central condensation of the star. This may be done if, for example, a heavier material like helium is deposited on top of the hydrogen-rich surface of the original star. Helium will compress the hydrogen gas, thereby shrinking the stellar radius.

Figure 3 shows how such a deposition may occur if the star exists in a close-binary system where the companion star, assumed to be initially the more massive, evolves first, converting hydrogen into helium in its core and expanding its radius in a normal way until its material overflows onto the secondary star. The secondary thereby accretes tremendous quantities of hydrogen and eventually some of the helium that was produced inside the former core of the primary. We might hypothesize that, if the secondary becomes pulsationally unstable in this manner, then we are witnessing a Beta Cephei star, one of a class of blue variable stars with periods of the order of a few hours. Note that these stars (triangles in Figure 2) do not have to have extremely high masses but can conceivably range as low as 6 $M_{\odot}$.

One further interesting speculation is that if the primary star loses all of its hydrogen envelope and some of its helium core as well, it will become a pure helium star. According to Figure 2, if the remnant (squares) has a mass greater than about 15 $M_{\odot}$, it will also become unstable due to the first pulsational mechanism. Perhaps some of the stars in the low-luminosity group of Wolf-Rayet stars may be accounted for in this way.

In conclusion, if these suggestions are correct, a unified interpretation can explain at least some of the members of the two most important classes of blue variable stars. Moreover, we have further important evidence for the existence of nuclear reactions in stars through their powerful destabilizing effect.

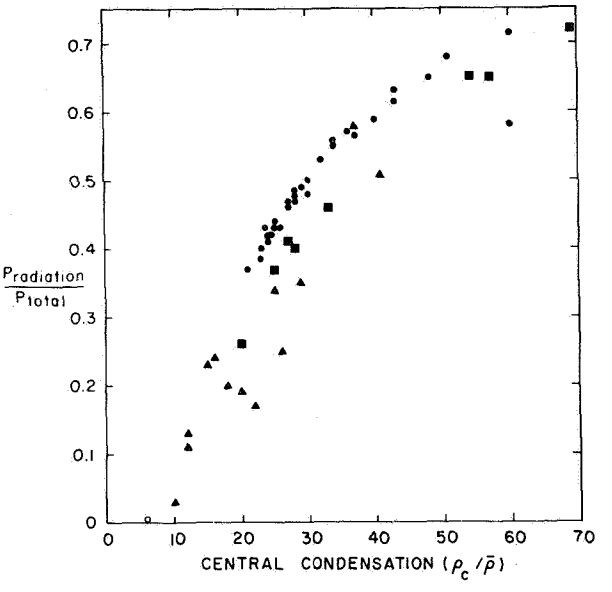


Figure 1

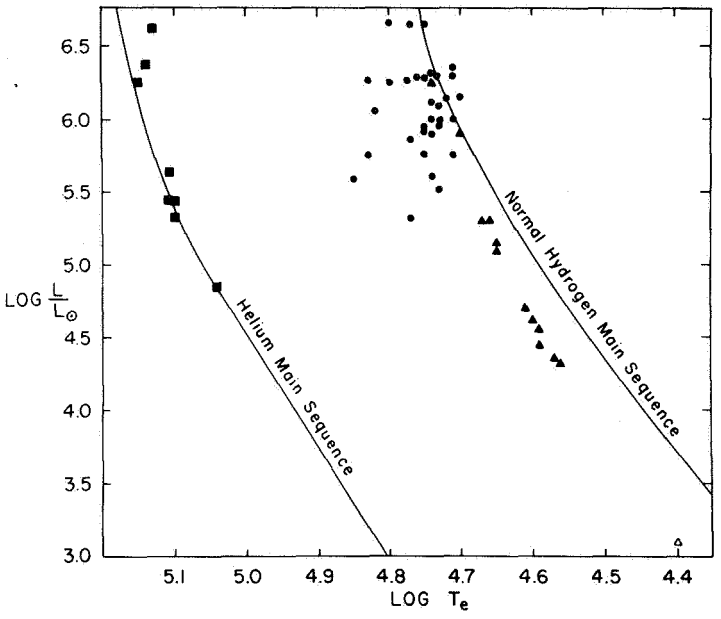


Figure 2

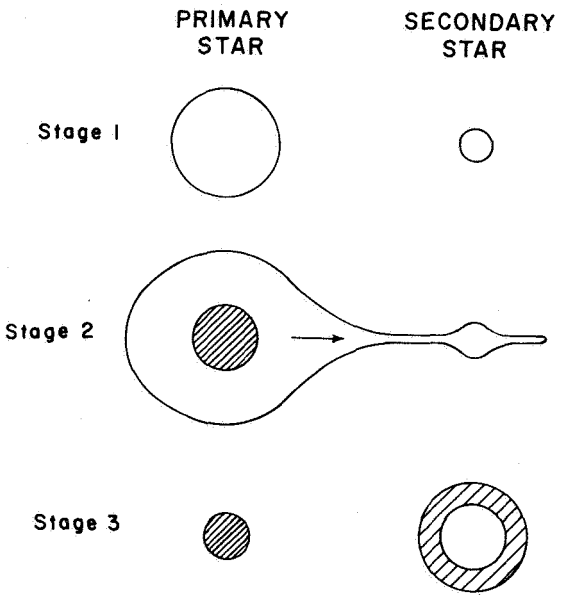


Figure 3

A HIGHLY EVOLVED STAR: SIGMA ORIONIS E

Daniel A. Klingelsmith

DR. KLINGELSMITH:

Considering the hour of the afternoon I imagine all of us are quite tired, and I would like to tell you about a star who is also quite tired. Sigma Orionis E is an old star, I think. It is the fifth and faintest member of a multiple star system in the constellation of Orion.

It is a helium star, and helium stars are characterized by strong helium absorption lines and weak hydrogen absorption lines—they are thought to be highly evolved objects.

Figure 1 shows the spectrum of Sigma Orionis E along with three of the other components in the multiple star system.

The wavelengths marked with dashed and dotted lines are the wavelengths of the helium absorption lines, and you can see that in Sigma Orionis E they are stronger than in the other stars that have similar spectral types.

The hydrogen lines, if you look carefully, are smaller—in the actual tracings they are quite small compared with the normal stars.

These plates are Coude spectrograms taken by Dr. Greenstein at Mount Palomar several years ago. They have been lent to Dr. Kurt Hunger and me, and we have done the abundance and atmospheric analysis of Sigma Orionis E.

We used data that covered the spectral range 3700 Å to 4700 Å, and in that region we measured the equivalent widths of 39 absorption lines, which are a measure of the number of atoms in the atmosphere. These were compared with a grid of theoretically computed equivalent widths from flux-corrected model atmospheres in which we varied the effective temperature, the surface gravity, and the chemical composition.

Table 1 lists the results of that analysis. This is what we think is the best fitting atmosphere to the star Sigma Orionis E: the effective temperature is in the neighborhood of 23,500° K; the log of the surface gravity is 4.05; the chemical composition by mass of hydrogen is 28 percent; and helium is 67 percent. All the other elements, which in this case happen to be only four, account for about 5 percent.

Table 1

Final Model Sigma Orionis E

Characteristic	Model	Normal Star
T_E	23,500° K ± 1000° K	–
Log g	4.05 ± 0.20	–
H	28%	70%
He	67%	27%
Z	5%	3%

This is to be compared with the normal unevolved abundances: 70 percent hydrogen, 27 percent helium, and 3 percent metals.

From this analysis we would say that indeed Sigma Orionis E has burned its hydrogen and converted it into helium. Now, since we know the effective temperature and the surface gravity of the star, and the distance to the star, we can determine the absolute luminosity, and from that we are able to get a measure, or an estimate, of the mass and the radius of the star.

The mass turns out to be 10 solar masses and the radius 5 solar radii, and this, if we are correct, is another indication of the fact that Sigma Orionis E is evolved.

Sigma Orionis A and B, whose spectrum is shown in Figure 1, is a close binary, and from the orbit calculations it has been determined that the total mass of the system A-B is about 50 to 100 solar masses.

It is a basic tenet of stellar evolution that the more massive star evolves faster, and yet the evolved star, Sigma Orionis E, is only a 10-solar-mass star. What has happened to Sigma Orionis E is that it has shed a large fraction of its mass by processes I neither know nor care to talk about today.

But here we see this tired old star who has thrown off approximately two-thirds of his original mass and is now just relaxing and going to enjoy himself.

CHAIRMAN:

Dr. Klingelsmith is now going to sit down and enjoy himself.

DR. KLINGELSMITH:

Yes.

MEMBER OF THE AUDIENCE:

How old is the star?

DR. KLINGELSMITH:

I was afraid somebody was going to ask me that question. I do not want to answer that. It is in Orion so it cannot be too old.

MEMBER OF THE AUDIENCE:

How long can it continue to burn that hydrogen before some other mechanism comes into play?

DR. KLINGELSMITH:

I would hate to speculate on where it is exactly in the course of its evolution, but I would imagine it is probably through burning hydrogen and what we are seeing is the result of hydrogen burning—a helium-rich envelope. The hydrogen has been burned up and we are burning helium or heavier elements in the core.

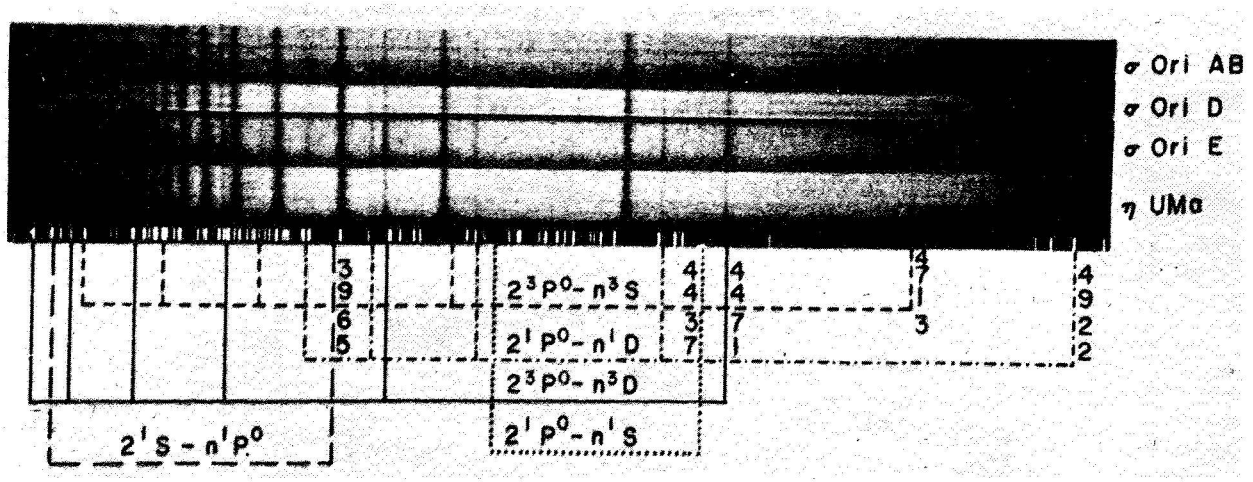


Figure 1 – Spectra of stars in σ Ori system, with η UMa for comparison. Helium series identified by different types of dashed and dotted lines.

COSMIC X-RAY LINE EMISSION

Stephen S. Holt

CHAIRMAN:

Next, Dr. Stephen Holt will discuss cosmic X-ray line emission.

DR. HOLT:

One of the primary goals of X-ray astronomy is the interpretation of the X-ray emission in terms of specific radiation mechanisms. We can get a crude indication of the dominant process in a discrete source from the overall shape of the spectrum. In other words, if the emission is nonthermal we should get something that looks like a power law, but if the overall spectral shape is better represented by an exponential, we have a good indication that thermal processes dominate the emission.

Unfortunately, these cannot be firm identifications because it is relatively easy to construct thermal models that masquerade as power laws and nonthermal models that masquerade as exponentials, at least over limited energy ranges. What we really need is a signature of the emission process. For example, strong differential polarization could be used as a signature of synchrotron emission, and, by the same token, the presence of K X-ray line emission in the spectra could be used as a positive indicator of thermal bremsstrahlung, since noncollisional processes do not give rise to such line emission.

About a year ago, we began to search for line emission from Sco X-1, the strongest X-ray source in the sky and one that has an overall spectral shape suggestive of a thermal origin. We concluded that with presently available techniques, namely, pointed Aerobee rockets and proportional counters, we could provide a positive indication of thermal emission if iron was more abundant in the source than the universal abundance. Since iron is always superabundant relative to its neighbors in the periodic table, because of its high nuclear stability, the lines from the highest ionization states of iron should be easily resolvable from the free-free continuum with the resolution available in proportional counters.

In Figure 1, the percentage excess above the continuum in a proportional counter resolution element centered about the iron lines is displayed as a function of temperature for universal abundance and for abundances which should be characteristic of both low and high mass supernova shells.

We expect to get statistics of 2% or better in such a resolution element; therefore we can certainly make a definitive statement about the supernova models because they give us more than a 20% excess in the vicinity of the Sco X-1 temperature, 60 million degrees Kelvin. We might even be able to say something statistically significant about universal abundance, which should have an excess of 5% or so.

We performed the experiment this year, and using a chi-square test on all of the data between 4 and 15 keV, we computed the best fit exponential folded forward through the detector response so that we could compare the data points directly with this best fit for fluctuations. In the upper trace of Figure 2, we have plotted the percentage difference between the data points and the best fit. The points are the 1/4-keV data, and the bins are 3/4-keV averages which are roughly consistent with the resolution obtained with proportional counters. Notice that the bump is at just the place we expected to see it, and its statistical significance is better than three sigma.

The next step was to add some iron-line emission to the trial exponential. The amount we put in was 0.5% of the total power in the 2- to 8-Å band, which is about what one would expect from universal abundance. We performed the same sort of analysis with this new input spectrum. As one can see in the bottom trace, this procedure makes the excess disappear.

So, on the basis of these data, I think we can certainly exclude the possibility of thermal emission from optically thin models which have much more iron than the universal abundance because they would give rise to much larger excesses than we observe. The inferred amount of iron-line emission, by the way, is about an order of magnitude below previously obtained upper limits, but upper limits are sort of negative results. We would like to ascribe positive significance to this if we can, and in order to do that we have to recall that a statistical significance of more than three sigma means that the odds are about a hundred to one that the bump is not just a statistical fluke. The odds are the same for any one of the 3/4-keV bins, but the only one that is high is, in fact, the one we expected to be high, and it has just the magnitude expected for universal abundance. It would seem that these data go a long way toward firmly establishing a thermal origin for the X-ray emission from Sco X-1.

MEMBER OF THE AUDIENCE:

Is there any distinction in your model between the emission from iron Kα and the emission from iron XXV?

DR. HOLT:

The two lines shown in Figure 2 are K-emission from XXV and XXVI. In other words, at about 60 million degrees Kelvin, there is nothing left in lower ionization states.

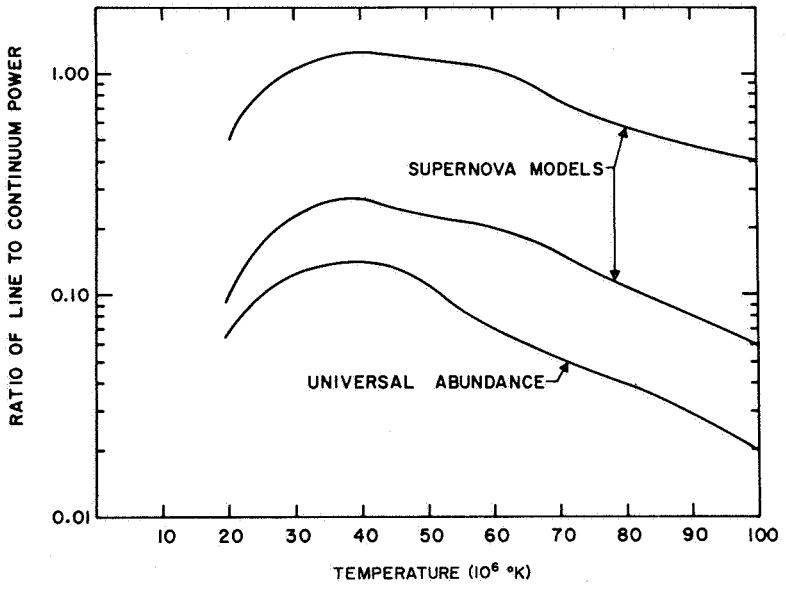


Figure 1

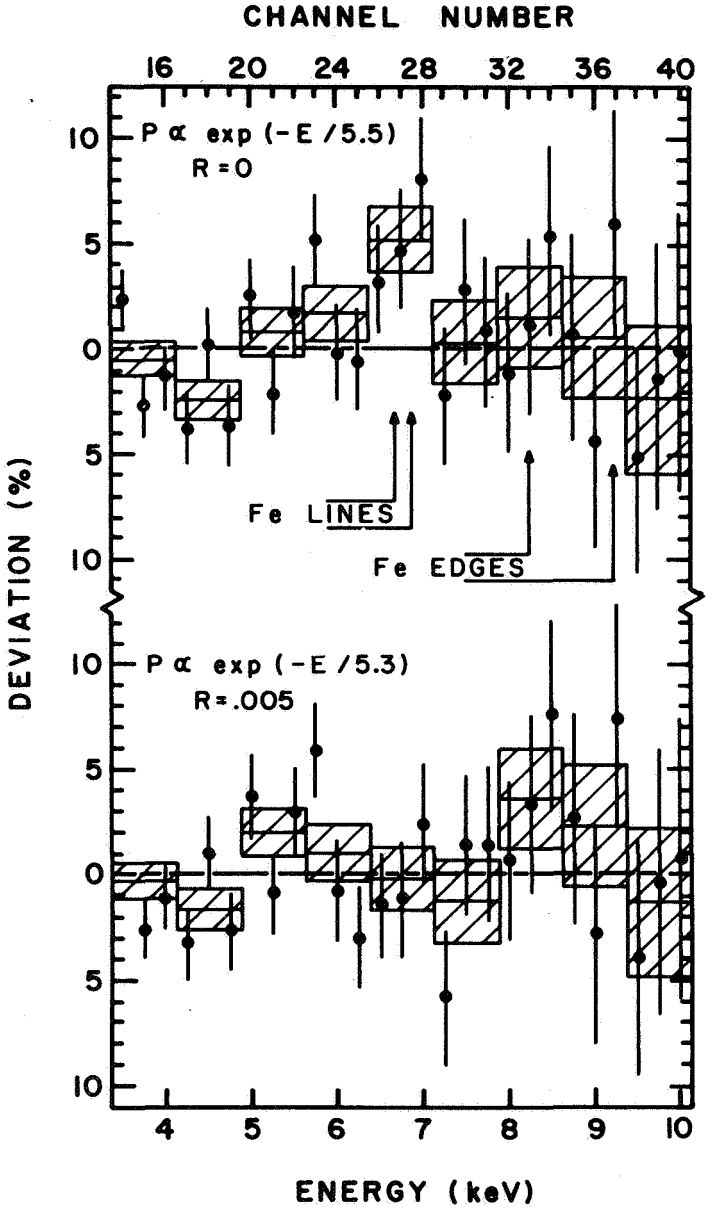

CHANNEL NUMBER
16 20 24 28 32 36 40
P α exp (-E /5.5)
R = 0
Fe LINES
Fe EDGES
P α exp (-E /5.3)
R = .005
DEVIATION (%)
10 5 0 5 10
4 5 6 7 8 9 10
ENERGY (keV)

Figure 2

THEORY OF PULSARS

Hong-Yee Chiu

CHAIRMAN:

I will call now for the paper, "Theory of Pulsars," to be given by Dr. Hong-Yee Chiu.

DR. CHIU:

Pulsars are objects that emit bursts of radiation chiefly in the 100- to 1000-MHz band at regular, periodic intervals ranging from 30 msec to several seconds.

The period is increasing so that it doubles about every million years. This increase in period shows that the timekeeping mechanism in pulsars must be rotation and cannot be vibration because a vibration period cannot change except for minute decreases.

In the rotation model, the observed optical emission and radio pulsed emission are the result of a very narrow beam corotating with the star so that it produces a pulse effect. This is like the beam from a scanning radar or the beacon of a lighthouse.

Now, the rotation model leads to severe difficulties. Centrifugal force prevents a star of a given radius from rotating faster than a certain period; otherwise, the star would fly apart. From the shortest period observed, 30 msec, the radius of a pulsar is inferred to be less than 100 km, and the only candidates are neutron stars which have a diameter of about 10 km. The density of a neutron star is about 10^{14} gm/cm^3, and it is thought to be the remnant of a supernova or exploding star.

Figure 1 shows a series of pictures of the pulsar NPO532 in the Crab Nebula, the only pulsar with detectable optical emission. It is the remnant of a star that exploded in 1054 A.D.; therefore its age is known to be exactly 915 years. These pictures were taken with an electronic camera, the GSFC Pulsar Hunter, on the 84-in. telescope at Kitt Peak by Steve Maran, Roger Lynds, and myself. Relative magnitudes, deduced from the photographs, are shown at the bottom of the figure. These 12 photographs show the entire pulsing cycle in 10 sequenced phases plus two additional phases that are a repetition of the first two phases. The time resolution from one picture to the next is about 3 msec. It is clearly seen that the star's brightness changes abruptly. Between the eighth and ninth photographs,

the star bursts into full brightness in less than 3 msec, faster than an incandescent bulb can be activated.

The total emissive power can be calculated from the known distance to the Crab Nebula, and a surface flux of 10^{17} erg/cm^2-sec—in radiowaves alone—can be inferred from the surface area of the neutron star. By comparison, all the electric power produced on Earth is equivalent to the radio emission from 1 cm^2 on the surface of this pulsar. If the emission were to come from a blackbody, then the temperature would have to be 10^{24} °K, a sheer impossibility.

Therefore, it is deduced that such a large brightness temperature can only arise from some coherent process such as laser emission or from some kind of ingenious device such as a klystron, but a klystron can only produce waves of a single frequency, and therefore, we can only deduce that the mechanism is a laser process.

In order to have a laser, you need two things: a proper energy level and a coherent energy source. The coherent energy source is the rotational energy of the neutron star, which is sufficient to keep the star shining for millions of years. Rotational energy is a pure form of energy, and its application and conversion are not limited by thermodynamic laws. This is the chief reason why all thermodynamic devices convert thermal energy into rotating shaft energy.

However, rotational energy alone is not enough to make a pulsar. Some mechanism must exist to convert the rotational energy into a coherent or laser emission. To achieve this, a magnetic field is needed.

As a star collapses into a neutron star, its magnetic field also collapses, and the field strength increases proportionally with the inverse square of the radius. Since the radius of the star changes by a factor of 10^5 during collapse, a modest initial field of 10^3 gauss will become an intense field of 10^{13} gauss. From our theoretical studies, we find that electrons in such a strong field behave very differently from electrons in weaker fields.

In weak magnetic fields, electron trajectories will be helices, and this motion can be decomposed into a circular motion perpendicular to the field and a linear motion parallel to the field. The radius of the circular orbit decreases as the field increases. As the magnetic field becomes stronger, the radius of the circular orbit approaches atomic dimensions. Then, the same process leading to quantization of atomic orbits must also operate, resulting in discrete energy levels in the circular motion. Orbits that are similar to atomic orbits result, but this quantization differs from atomic

quantization in that it produces energy levels in the perpendicular motion only and leaves the energy of motion parallel to the field unchanged (Figure 2). So, we have a very special kind of atom, where electrons possess only one degree of freedom with discrete energy levels.

The discreteness of the energy levels makes possible a stellar laser, and the continuous energy distribution along the field lines allows it to be a broadband laser. It turns out that the laser process is so efficient that, when it operates on the surface of a neutron star, saturation is achieved. That is, the surface is emitting at the maximum rate allowed by electrodynamic theory. This explains why all pulsars emitting in the same frequency regime have approximately the same intensity, independent of their period.

This laser is excited by an electric field of the order of 10^{10} V/cm that is produced by the rotation of the neutron star's magnetic field. The electric field also acts as a coupler to the rotational energy of the star. The emerged radiation is guided by magnetic field lines into very sharply defined beams (Figure 3). Then, as these beams rotate, they pass by us, producing a pulsing effect. That is why the star can be "turned on" in such a short time. This model in its simplest form is already able to explain a number of observed phenomena such as the polarization changes.

MEMBER OF THE AUDIENCE:

Can you explain why the pulses do not have constant amplitudes but fluctuate with time?

DR. CHIU:

Not yet. I think this type of problem belongs to what we should call the secondary phenomena of astrophysics. For example, to explain the radiation properties of stars, first we explain the basic mechanism of radiation. This we know to be blackbody radiation. But you can ask, does this explain the variable star? It turns out that variable stars also radiate via the blackbody radiation mechanism, but there is something inside the star causing a change in its radius, and this change in radius produces the variation in light intensity. I think, at the present moment, we should be chiefly concerned with the high brightness temperature. Other effects cannot be studied unless one understands the basic mechanism of radiation.

MEMBER OF THE AUDIENCE:

Would you not expect a fixed polarization with that model? One of the interesting observations is that the polarization of radiation from some pulsars changes with time periodically.

DR. CHIU:

In the present theory, radiation comes from the surface of the star, and at any instant, we are looking toward the star at a particular angle. As this medium is embedded in a strong magnetic field, it is highly doubly refractive. We know that light, transmitted through a double-refractive medium (such as calcite), is split into two components, the ordinary component and the extra-ordinary component. It turns out that, in a dense plasma in a strong magnetic field, the only linear polarization allowed is the one which is perpendicular to the field. Of course, the degree and direction of polarization changes with both the angle of emergence and the field. Since we view the radiation at all possible angles of emergence when we scan the surface for one entire period, we naturally expect the polarization to change with phase. At present, our theory can explain qualitatively the change of polarization, but as we dig more deeply into our theory for more quantitative information, we expect that a more detailed comparison with observation will be made.

MEMBER OF THE AUDIENCE:

How did you obtain population inversion in the laser process so that you have coherent emission?

DR. CHIU:

This turns out to be quite involved. In fact, there are several incorrect papers in the literature on this subject, including some of my previous papers. However, the condition for population inversion in an ordinary laser operating between two discrete energy levels is well known. A population inversion, or negative temperature, is achieved when the number of particles in the higher energy level exceeds the number of particles in the lower energy level. One can show that, in a discrete two-level system, merely an excess of particles in the upper energy level automatically gives rise to laser emission.

In the case of continuum energy levels, such as that which we have been discussing, the criterion for laser emission is modified. Since the distribution of energy is continuous, population inversion can be defined in a broader sense: There are more electrons with higher energy than electrons with lower energy. This population inversion cannot go on forever because in a finite system the electron distribution must eventually decrease with energy, otherwise the total number of electrons will become infinite. Therefore, in a system with a continuous energy distribution, population inversion (negative temperature) is always accompanied by normal population (positive temperature). In order for laser action to take place, the amplification must be stronger in the parts of the spectrum having negative temperatures.

The criterion for laser emission for a displaced Maxwellian distribution (a Maxwellian distribution displaced by a coherent motion in a particular direction), under certain assumptions, is

$$\int f(p) \frac{\partial}{\partial p} \frac{w(p)}{p} \, dp < 0,$$

where $f(p)$ is the distribution function of momentum p, and $w(p)$ is the emissivity. We found that, although the form of $w(p)$ for normal electrons does not satisfy this criterion, $w(p)$ for electrons in a strong magnetic field does because of the one-dimensional character of the orbits. Therefore, laser action can take place with broadband emission.

MEMBER OF THE AUDIENCE:

Would you predict monochromatic radiation for this laser process?

DR. CHIU:

No, it is not monochromatic. The radiation is produced by a bremsstrahlung process which yields a continuum spectrum. It is the one-dimensional bremsstrahlung process that gives rise to the laser process, hence the spectrum is a continuum. The shape is proportional to $\nu^{-\alpha}$ where ν is the frequency and α is approximately 2.

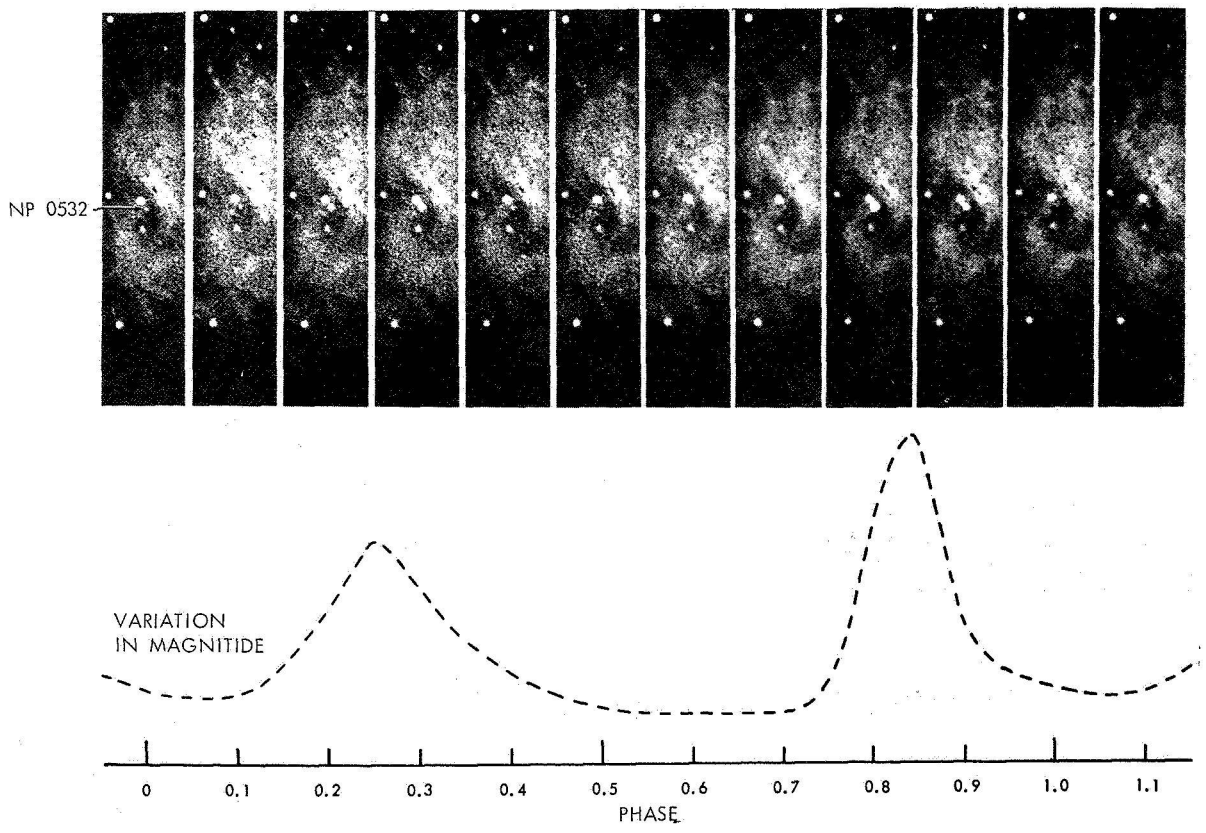

NP 0532
VARIATION
IN MAGNITIDE
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
1.1
PHASE

Figure 1

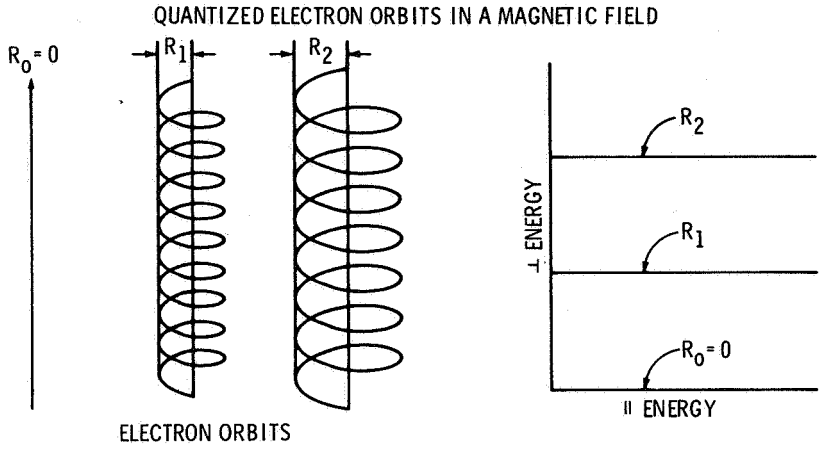


Figure 2

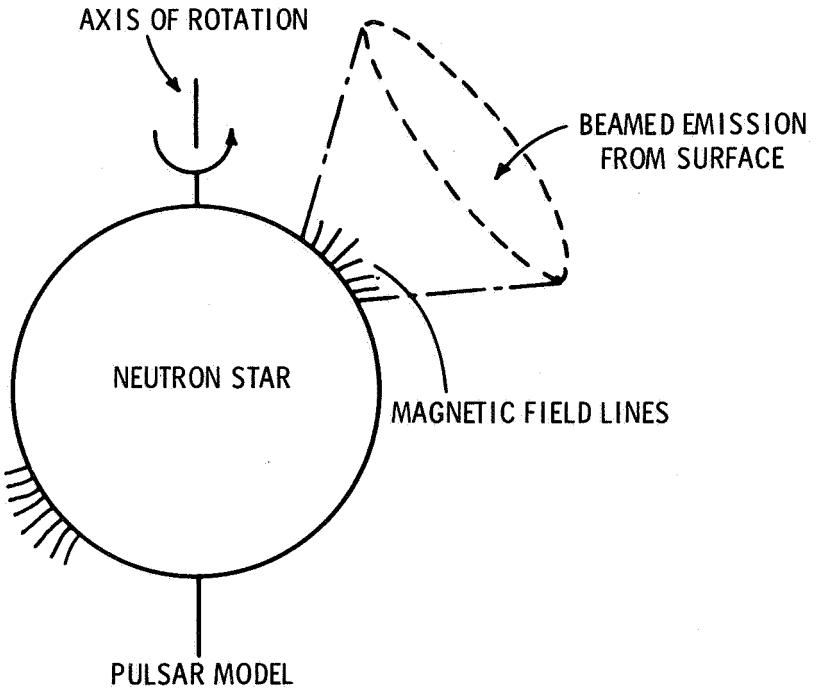


Figure 3

II
TECHNOLOGY

Page intentionally left blank

OPENING REMARKS BY DR. JOHN F. CLARK DIRECTOR OF GODDARD SPACE FLIGHT CENTER

This is part of our second annual review in science and technology. Although it is not quite fair to the technology part of these two days, we are really borrowing a very successful format that we first used 11 months ago when we had our initial experiment in a series of five-minute talks, summarizing some of the most significant and exciting results of the year in various areas of science.

After considering the format of our annual ART/SRT review, we decided this year to merge the two occasions and to expand the format to include all areas of GSFC's activity, not only in supporting and advanced research and technology but in the results of our space flight projects as well. After all, in many cases, earlier supporting research and technology made possible these results. At the same time, these results made it evident that new experiments should be undertaken and that new items of technology should be developed.

So, recognizing the essential integrity of the entire program, we felt that the best method of presentation was to bring the entire program together as an integrated whole.

It is now my pleasure to introduce the chairman for today's session, our Deputy Director of Technology, Mr. Robert Mackey.

REMARKS BY MR. ROBERT J. MACKEY, JR., CHAIRMAN

Thank you, Dr. Clark. I would like to wish you all good morning and welcome you to the second day of the science and technology review.

Today we are going to hear a number of papers dealing with technology accomplishments at Goddard Space Flight Center in the last year. As Dr. Clark indicated, the format for this technology presentation has been adapted from that which was successful in the science area.

In order to keep on schedule, I would like to limit questions from the floor to one or two short ones, if possible.

STUDIES OF ORBIT LIFETIMES OF VENUS AND MARS ORBITERS

Robert E. Coady

CHAIRMAN:

The first speaker will be Mr. Robert Coady, who will discuss orbit lifetimes of Venus and Mars orbiters.

MR. COADY:

In planning a mission to another planet, it may become necessary to select an orbit around the planet that would have a very low pericenter and a high apocenter. Such satellite orbits are of interest when it is desired to study large volumes of space with a maximum scientific payload. The orbital lifetime of such a satellite is very sensitive to perturbations since the pericenter is just above the atmosphere of the planet. Not only is the orbit affected by the oblateness of the planet, but it is also disturbed by other planets and possibly by the atmosphere of the central planet.

For such a mission, it is important to evaluate not only the coupling between the oblateness effect and the long-term motion caused by other planets but also the effect of the periodic, atmospheric perturbations. Either of these can cause the pericenter radius to decrease by an amount that will result in the satellite's impacting the planet. Thus, the orbit must be chosen to allow a sufficient lifetime for conducting the required scientific experiments and to meet the stringent quarantine conditions that arose from an international agreement not to disturb the ecology of the planets Venus and Mars prior to 1989. The latter necessitates long-term orbit prediction based on the initial conditions of the orbit after the retrofire insertion maneuver.

From a scientific point of view, the time history of the shape and orientation of the orbit is important, since the type and amount of scientific information to be obtained depends markedly on the orbit geometry. Apart from the lifetime problem, it is advisable to have an accurate time history of the pericenter radius as well as the histories of the other orbit parameters. Additionally, for preliminary mission design, orbit evolution information should be available for a wide variety of initial (planetary orbit insertion) conditions. The pericenter-radius history can be accurately predicted from the time history of the orbit's eccentricity and semimajor axis.

The major purposes of the study reported here are as follows:

1. To develop techniques for predicting the lifetimes of planetary satellite orbits–primarily Venus and Mars satellite orbits.

2. To evaluate the effects of the major perturbations acting on the spacecraft orbit: the oblateness of the planet, third-body perturbations, and atmospheric drag.

3. To recommend the software development needed for preflight mission planning and in-flight determination of orbit insertion conditions.

It is of interest not only to determine and assess the effects of the major perturbations acting on the satellite but also to evaluate or develop, in the context of accuracy and efficiency, approximation techniques available to us for orbit prediction. In addition, an expected output of the study is a recommended hierarchy of computer programs–each level using different mathematical techniques–through which it will be possible to resolve the orbit prediction problem in actual mission designs and flight control situations.

There are three principal disturbing effects that will cause satellites of either planet to deviate from two-body motion. The three effects are third-body perturbations, central-planet nonsphericity, and atmospheric drag. The major disturbing forces acting upon Venus orbiters are the third-body (i.e., solar) effects and, below about 300 km altitude, the Venusian atmosphere. The sharp drop in density found in the most recent models of Venus' atmosphere and the very low estimates of the Venusian oblateness make orbit prediction relatively easy.

For Mars, the story is quite different. Even though the surface density of the Martian atmosphere is much lower than that of Venus' or Earth's, the maximum density models for Mars predict significant atmospheric effects on satellite lifetime at very high altitudes. The possibly very dense upper Martian atmosphere is not the only difficulty. The oblateness coefficient J_2 for Mars is nearly twice that of the Earth. The very complex coupling between oblateness and third-body perturbations makes it reasonable to expect strong resonance effects between oblateness and medium-periodic, third-body effects, and such resonance could result in very short lifetimes. Figure 1 shows an example of strong coupling with third-body perturbations dominant. Here, even though the oblateness effects are small, they are still large enough to prevent impact. The oblateness could just as easily be the cause of impact in some situations where third-body effects alone would give long lifetimes. The solid line in Figure

1 shows the result of numerical integration of the actual equations of motion (Cowell algorithm, double precision). This can be compared with a solution obtained with an approximation technique (circles). The solutions are almost identical in the parameter that is shown (pericenter radius). The numerical integration required about two hours of computer time, but the approximation technique gave similar results in less than one minute of computer time.

The effects of atmospheric drag on orbit evolution are far more complex than the other perturbations. The difficulties become almost insurmountable at high altitudes because of the dynamic nature of the upper atmosphere. Fortunately, some simplifying assumptions permit us to study the drag effects in several atmospheric models. The primary assumption is that of a static density profile where the density is a function of altitude only. By studying the effects of a worst-possible density function, we can draw some conclusions about the ranges of orbit altitudes where atmospheric effects may or may not be ignored.

The technique used to study the drag problem is a numerical integration of a set of numerically averaged differential equations. (Atmospheric rotation is neglected.) The technique does not suffer from convergence problems at high eccentricities, and no assumptions are made about the density-versus-altitude function. If we keep in mind that the results obtained by this technique are not the complete solution to the atmosphere problem, we can gain a valuable feeling for the dynamics and the numbers.

Figure 2 shows the effect of drag on Martian-satellite lifetime. The dependence of lifetime on apocenter altitude, for a given pericenter altitude, is clear from examination of these curves. These data are based on the maximum-density model atmosphere for Mars and a ballistic coefficient for a Planetary Explorer spacecraft.

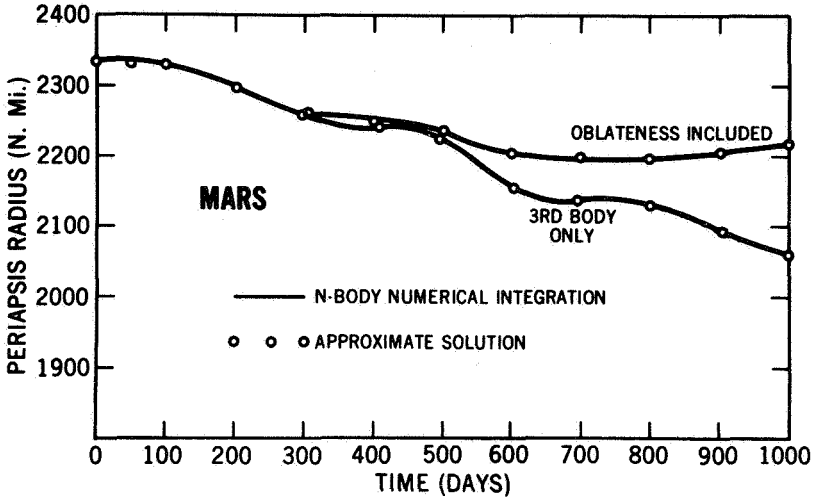


Figure 1

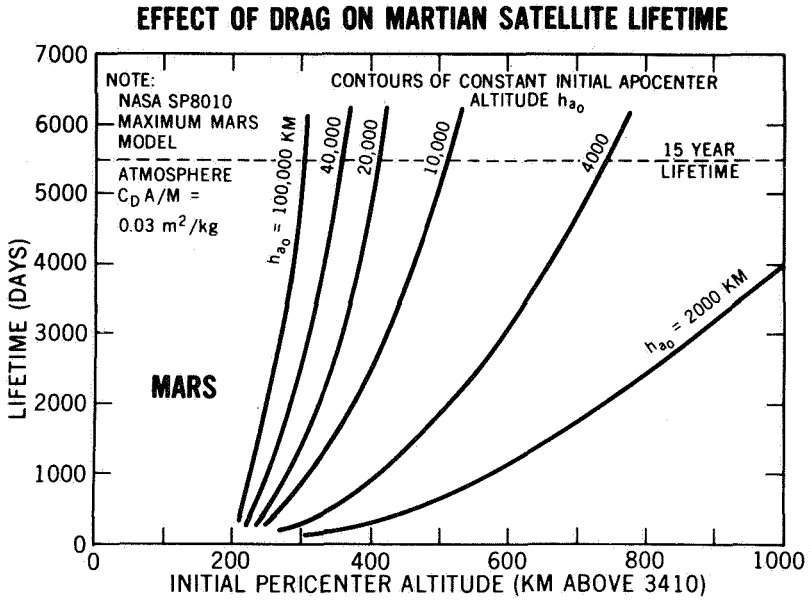


Figure 2

GENERALIZED MULTISTEP METHOD AS APPLIED TO TRAJECTORY DETERMINATION

Carmelo E. Velez

CHAIRMAN:

Our next speaker is Mr. Carmelo Velez, who will speak on the generalized multistep method of numerical integration for trajectory and orbit determinations.

MR. VELEZ:

The objective of this study is to develop new methods of numerical integration, with the goal of reducing the computer time required by our present orbit determination and parameter estimation systems.

It is well known that a large portion of the total machine time used by all the orbit determination systems, lifetime study, and physical parameter estimation programs, such as those used for geopotential or lunar potential estimation from tracking data, is spent numerically integrating a system of differential equations. It is for this reason that any new, more efficient integration algorithm, when utilized in one or more of these systems, can result in considerable savings in the amount of computer time used. Such a goal has already been partially realized (Reference 1) by the incorporation of special methods of local error control in our current Cowell orbit determination system, and it is felt that even more efficient methods can be designed and implemented.

One of the areas of current research that appears promising for this purpose is the development of new methods based on high accuracy numerical integration formulas known as generalized multistep. These new methods are being designed specifically for the numerical integration of the differential equations of the type that occur in our orbit determination programs (Reference 2) and are patterned after the Gauss-Jackson or Cowell type of algorithm, which has been the standard numerical integration method here for many years.

These new generalized multistep formulas can be characterized as stable numerical integration formulas which are more accurate than the currently available methods while using the same number of backpoints. The basic idea behind their development is the following:

There is a well-known theorem of numerical analysis (Reference 3) which states that for a K-step method to be stable its order cannot exceed $K + 2$. Here, a K-step method is a numerical integration formula that uses K equally spaced backpoints to integrate a differential equation, stability is a property of the method that insures reasonable error propagation, and the order of the method is a measure of its accuracy (so, high order methods are desirable). This stability result constrains the order of accuracy of our currently used multistep methods to $K + 2$. The new generalized multistep methods, however, circumvent this stability restriction and achieve orders of accuracy as high as $2K + 1$ while maintaining stability (References 2, 4, and 5).

These highly accurate formulas may lead to more efficient computer algorithms by allowing the use of large step sizes while maintaining the accuracy of the current methods with smaller step sizes. This means that whenever a minimum accuracy requirement on the computation of a trajectory is specified, generally by the accuracy of the measurements, the new formulas could achieve this accuracy with a smaller number of arithmetic operations, thus saving computer time. Experimentation to date has been encouraging, but more work is needed.

In the following, the results of some preliminary accuracy comparisons between the new generalized methods and the classical Gauss-Jackson method are presented. These tests were performed by Dyer and were presented at the 1969 AIAA Conference (Reference 2).

Figure 1 displays the approximate number of digits of accuracy maintained while computing a 900-minute satellite trajectory, using methods based on the Gauss-Jackson formula and on the new generalized multistep formulas. This trajectory represents roughly 8 or 9 revolutions of a low-ecentricity Earth satellite. In this test, the generalized methods using 4 and 5 points were compared with a Gauss-Jackson using 7 points, with step sizes ranging from 1 to 4 minutes.

The accuracy advantage of the higher order generalized multistep methods is clear, the 5-point generalized formula performing better than, and even the 4-point generalized comparing favorably with, the 7-point Gauss-Jackson. One might extrapolate from this diagram the fact that to achieve 7 digits of accuracy, or approximately 1 m, the Gauss-Jackson required a 2-minute step size, but the generalized 5 yielded this accuracy with a 3-minute step size, thereby indicating possible savings in computer usage with the generalized 5 method.

Figure 2 shows the results of an integration of a 5-day trajectory with both the generalized method and the Gauss-Jackson using the same number of points and a 4-minute step size. This test shows that the new highly accurate methods are also stable, since the error propagation over the 5-day period is reasonably smooth, whereas an unstable method would have displayed an exponentially increasing error. Also, as in Figure 1, one might determine from extrapolation that to achieve the 6- or 7-place accuracy of the generalized method, the classical method would require a much smaller step size than the 4 minutes used, again indicating an advantage in efficiency for the generalized method.

The current effort will include further development of generalized methods and further comparisons with existing procedures, such as Cowell and power series. These tests will be performed on various orbit types and will include formal efficiency as well as accuracy comparisons. Those methods that are shown to offer distinct advantages in efficiency will be incorporated in our orbit determination and parameter estimation systems.

CHAIRMAN:

Are there any comments or questions for Mr. Velez?

MEMBER OF THE AUDIENCE:

Is there any explanation why you have some improvement in accuracy even though the step size increases?

MR. VELEZ:

Yes. It is roughly the same as trying to use a very-high-order interpolating polynomial and a very small step size. It generally happens that the round-off error deteriorates the answer. And as you increase the step size, the accuracy improves due to the decrease in this round-off error.

The point here is that these formulas are highly accurate—their order is very high. A 5-point formula is equivalent to an 11th-order method, so that the lower region of step sizes results in a dominance of round-off error.

REFERENCES

1. Velez, C., "Local-Error Control and Its Effects on the Optimization of Orbital Integration," NASA Technical Note D-4542, June 1968.

2. Dyer, J., "Generalized Multistep Methods in Orbit Computation," *J.A.C.M.* 15: 712-719, 1968, paper presented at the AIAA Aerospace Computer Systems Conference, Los Angeles, Calif., Sept. 8-10, 1969.

3. Dahlquist, G., "Convergence and Stability in the Numerical Integration of Ordinary Differential Equations," *J.A.C.M.* 12: 124-135, 1965.

4. Gragg, W. B., and Stetter, H. J., "Generalized Multistep Predictor-Corrector Methods," *J.A.C.M.* 11: 188-209, 1964.

5. Hansen, E., "Cyclic Composite Multistep Predictor-Corrector Methods," paper presented at the AIAA Aerospace Computer Systems Conference, Los Angeles, Calif., Sept. 8-10, 1969.

ACCURACY–STEP SIZE DIAGRAM

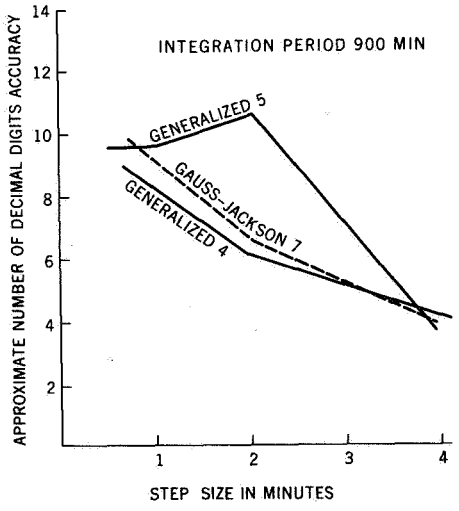


Figure 1

ACCURACY–INTEGRATION PERIOD DIAGRAM

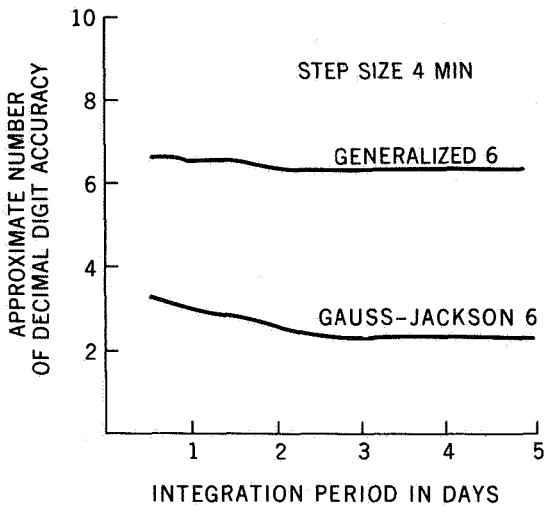


Figure 2

TESTS AND COMPARISONS OF DIFFERENT SETS OF GEOPOTENTIAL COEFFICIENTS FOR ORBIT DETERMINATION AND PREDICTION

James G. Marsh

CHAIRMAN:

The next presentation will be given by Mr. James Marsh. His subject is the comparisons of geopotential coefficient models and the observed orbital parameters of geodetic satellites.

MR. MARSH:

In order to fully utilize the high precision of present and future satellite tracking systems, we must pay particular attention to the forces acting on the spacecraft. Error analysis studies have shown that, for drag-free satellites, by far the largest errors in orbit determination are introduced by uncertainties in the gravity field. In this report I would like to present some recent results of definitive orbit determination for the Geodetic Earth Orbiting Satellites (GEOS 1 and GEOS 2) when different sets of geopotential coefficients are used.

GEOS 1 and GEOS 2 are geodetic satellites which were launched as part of the National Geodetic Satellite Program under the direction of the National Aeronautics and Space Administration. The orbits and spacecraft were chosen so that the perturbations due to air drag and solar radiation pressure would be small.

The orbital inclination for GEOS 1 is 59° prograde, and GEOS 2 was launched into a retrograde orbit with an inclination of 106°. GEOS 1 has an orbital frequency of nearly 12 revolutions per day, which gives rise to a resonance with the 12th-order gravity coefficients. GEOS 2 has a smaller orbital period, making it resonant with the 13th-order coefficients. The total effect of these resonances is about 500 m along track in each case.

For this study, orbits over approximately 6 days were computed using the precision reduced optical observations recorded by the Space Tracking and Data Acquisition Network Minitrack Optical Tracking System and the Smithsonian Astrophysical Observatory (SAO) Baker-Nunn cameras. The accuracy of the optical data is of the order of 2 seconds of arc, which amounts to a positional error of approximately 15 m for the GEOS satellites.

Orbital solutions were computed using eight unclassified sets of gravity coefficients. These eight sets of coefficients were published by labs,

agencies, and universities, such as the Naval Weapons Laboratory, Applied Physics Laboratory (APL), SAO, Ohio State University, and the University of California at Los Angeles.

The best GEOS 1 orbits were obtained using the SAO coefficients. The accuracy of these orbits was of the order of 25 to 30 m. Orbit errors as large as 500 m were observed, however, when some of the other models were used.

The GEOS 2 orbits were not as good as those obtained for GEOS 1 until the 13th-order resonant coefficients were derived by GSFC and APL.

Table 1 shows typical results for a 6-day arc of GEOS 2 data. The agreement between the computations and the observations was significantly improved when the models were modified by adding improved values for the 13th-order resonant coefficients.

Table 1
RMS's About Fitted Orbits Determined from GEOS 2 Optical Observations When Different Published Sets of Geopotential Coefficients are Used
[April 28–May 4, 1968 (1098 Observations)]

Model	Resonant Terms Added	Seconds of Arc
SAO M-1 – 1966		17.4
SAO M-1	GSFC and APL 13th order terms	3.1
APL 3.5 – 1965		59.4
APL 3.5	GSFC and APL 13th order terms	5.8
KAULA – 1967		16.7
KAULA	GSFC and APL 13th order terms	9.3

Currently, the modified SAO M-1 model is being used for GEOS 2 orbit computations.

Approximately 80% of the resonance effect is along track. Figure 1 shows the along-track position differences when some of the trajectories are compared with the modified SAO M-1 reference orbit. The reference orbit is good to within about 30 m. Position differences are as large as 1000 m. This is in the definitive period and is not a prediction.

In conclusion, the use of an accurate representation of the Earth's gravity field plays a very important role in definitive orbit determination. Current and future efforts at GSFC are expected to produce a more general and accurate set of gravity coefficients and station coordinates than is presently available.

CHAIRMAN:

Are there any comments or questions for Mr. Marsh?

MEMBER OF THE AUDIENCE:

Do you get similar numbers on crosstracking, or are they much smaller?

MR. MARSH:

They are much smaller, yes.

MEMBER OF THE AUDIENCE:

About 80% of the orbit error is along track, 20% is crosstrack and radial.

MR. MARSH:

Yes.

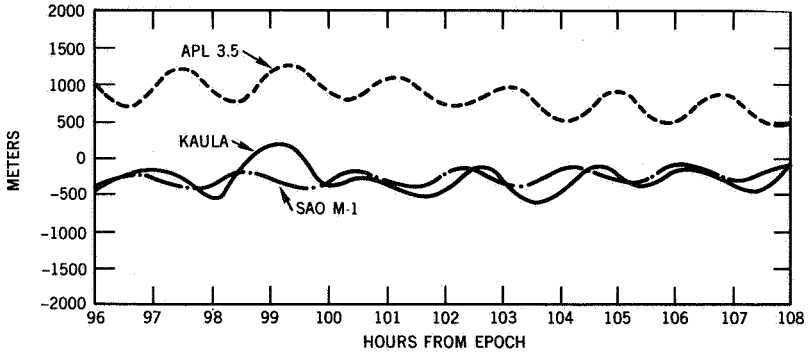


Figure 1–Along-track position differences on 5/2/68 obtained when 6-day GEOS 2 ephemerides covering 4/28/68 to 5/4/68, computed using different sets of geopotential coefficients, are compared. SAO M-1 with GSFC and APL 13th-order resonant terms versus SAO M-1, APL 3.5, and Kaula.

NEW RESULTS IN RESONANT SATELLITE GEODESY

Carl A. Wagner

CHAIRMAN:

Our next speaker is Mr. Carl Wagner, who will report some new results in the use of resonant satellite data in geodesy.

MR. WAGNER:

I would like to talk briefly about the work I am doing in the incorporation of deeply resonant satellite data into the solution for the geopotential. A resonant satellite is one whose mean motion is commensurate with the Earth's rotation rate. From the tracking standpoint, such satellites (of which there are now literally hundreds of examples) are important because they can suffer long-term oscillations along track of up to 180 degrees in amplitude, due to very small, but resonating, longitude-dependent forces in the Earth. From the geodetic standpoint, such satellites are important because these large, readily observable oscillations should enable us to determine (to great accuracy) these tuned resonating forces, which are otherwise hard to observe.

In the past few years geodetic solutions have incorporated data from many low-altitude satellites weakly resonant with high order and degree geopotential terms, for example, the two GEOS orbits. My own work has dealt with the synchronous communications satellites of high altitude which are strongly resonant with a large number of low-frequency terms in the geopotential.

Table 1 shows the deeply resonant satellites I was analyzing earlier this year, their mean motions, inclinations, and data arc lengths. Recently I have added ATS 1, 3, and 5 to this list.

Figure 1 shows the quality of determination of the spherical harmonics of gravity H_{lm} which can be expected from a combined solution with the data from the deeply resonant high-altitude satellites. It can be seen that about half of the spherical harmonics to order and degree 8 significantly affect this type of satellite. The l subscript in this diagram is an index of the latitude frequency of the term; the m subscript an index of its longitude frequency. S indicates strongly determined by this data, W indicates a weak dependence. For example, the (2, 2) harmonic, associated with equatorial ellipticity, is strongly determined by the tracking data from these satellites.

Table 1
Resonant Satellites Analyzed in Current Low-Order Gravity Solutions

Satellite	Orbit Frequency (rev/day)	Inclination (degrees)	Number of Arcs Being Used (100-500 days per arc)
SYNCOM 2	1	29-33	6
SYNCOM 3	1	0-3	4
Early Bird	1	0-3	2
Intelsat 2F-3	1	1	1
Intelsat 2F-1 (Looney Bird)	2	18	1
Cosmos 41	2	65-66	2
Cosmos 41 rocket	2	67-69	2
Molniya 1	2	64-65	4
Molniya 4	2	65	1
Molniya 7	2	65	1

To give an example of the effect of these terms, I have determined the orbit of Intelsat 2F-8 over a one-year period from GSFC Minitrack data, using nonresonant and resonant gravity fields. The results are shown in Table 2, from which it can be seen that the position error in these orbits can be reduced by over an order of magnitude with the use of a simple, and incidentally realistic, resonant gravity field which is quite close numerically to the nonresonant fields.

One-Year Orbit Determination for Intelsat 2F-1 (12-Hour Satellite)

Geopotential Field Used	Average Position Error (km)
Smithsonian 1969–'COSPAR' (nonresonant)	180
Smithsonian 1966–Standard Earth (nonresonant)	90
6-parameter resonant field	9

Clearly, considerable geopotential improvement can be achieved with constraints provided by the data from these deeply resonant satellites. At present I am reducing these constraints to a form suitable for incorporation into a full-field geopotential recovery program so as to utilize strongly resonant satellite data most efficiently.

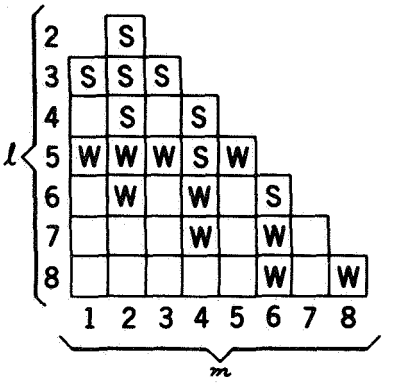


Figure 1—Quality of determination of H_{lm} from deeply resonant high-altitude satellites (S = strong, W = weak).

LUNAR GRAVITY FIELDS DETERMINED FROM APOLLO 8 TRACKING DATA

Theodore L. Felsentreger

CHAIRMAN:

Our next speaker is Mr. Ted Felsentreger, who will report on the determination of lunar gravity fields from Apollo 8 data.

MR. FELSENTREGER:

This effort represents an attempt to derive a lunar gravity model for an Apollo-type orbit that would have adequate orbit prediction capabilities for use in subsequent Apollo missions. The orbits of the first few Apollo missions can be characterized as having low altitudes and eccentricities, and relatively low inclinations.

The motivation for the effort extends back to the time of the Langley Lunar Orbiters, which were the first lunar orbiting spacecraft launched by the United States. It was noticed that the lunar gravity models used for the orbit determination of these satellites had rather poor orbit prediction capabilities and produced unexpectedly large range-rate residuals around perilune, the closest approach to the moon. In addition, gravity models derived from Lunar Orbiter tracking data seemed to have the same shortcomings. It was recognized that lack of an adequate model of the lunar gravity field could present a serious problem for the Apollo missions, especially the landing missions.

The flight of Apollo 8, the initial manned lunar mission, presented the first opportunity to analyze tracking data from an Apollo-type orbit for the effects of lunar gravity. Apollo 8 made 10 revolutions of the moon, the last eight of which were near-circular at an altitude of about 112 km and an inclination of about 12° retrograde relative to the lunar equator. We took range-rate (Doppler) data from these last eight orbits and derived a number of gravity models through degree and order 6 in spherical harmonics. I have chosen one of our fields for comparison with that adopted for the Apollo 10 and 11 missions. It should be mentioned here that, to the best of our knowledge, this was the first time that data from the Manned Space Flight Network (MSFN) were used in the derivation of lunar gravity models.

Table 1 presents the two models in unnormalized spherical harmonic coefficients. We derived Field 15 from orbits 6, 7, and 8 in a single-arc

solution for state (i.e., position and velocity) and harmonic coefficients. The R2 Field was determined by Boeing from long period variations in the orbital elements of some of the Lunar Orbiters and was adopted for the Apollo 10 and 11 missions.

Table 1
Lunar Gravity Fields

R2 Field				Field 15	
n	m	C_{nm}	S_{nm}	C_{nm}	S_{nm}
2	0	-2.07108×10^{-4}	0	-1.9880×10^{-4}	0
3	0	-0.21000×10^{-4}	0	-0.7567×10^{-4}	0
3	1	-0.34000×10^{-4}	0	0.5386×10^{-4}	-0.2908×10^{-4}
2	2	0.20716×10^{-4}	0	0.4909×10^{-4}	-0.1405×10^{-4}
4	4	0	0	0.0011×10^{-4}	-0.0020×10^{-4}

Figure 1 shows the range-rate residual patterns produced by the two models for orbit 8. Field 15 provided significantly smaller residuals than did the R2 Field, but this is not surprising in view of the fact that orbit 8 was used in the derivation of Field 15.

However, Figure 2 shows the range-rate residual patterns produced for orbit 5, which was not used in the derivation. Again, the residuals produced by Field 15 are smaller than those for the R2 Field. In addition, estimates were made of the in-track, cross-track, and perpendicular errors induced by the two models for orbit 5, and Field 15 consistently produced errors about 1/5 of those for the R2 Field. However, Field 15 failed to predict the eccentricity increase from orbits 3 through 10, while the R2 Field did predict this.

We also compared the two models using Apollo 11 data. The rms of the residuals over one orbit were about the same for each model; in addition, in this case, Field 15 did predict the eccentricity evolution.

Tests on Field 15, and other promising models, are continuing using Apollo data. We also hope to derive a more general gravity model from many of the lunar orbiting spacecraft–the Lunar Orbiters, Apollos, and

Anchored IMP. Also, we are implementing a computer program with the ability to solve for mascons (mass concentrations), in addition to spherical harmonic coefficients.

CHAIRMAN:

This concludes the papers on trajectory and orbit determination.

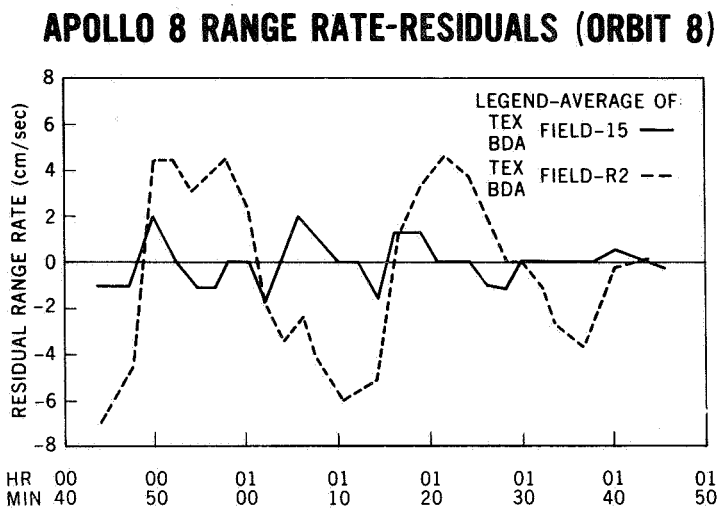


Figure 1

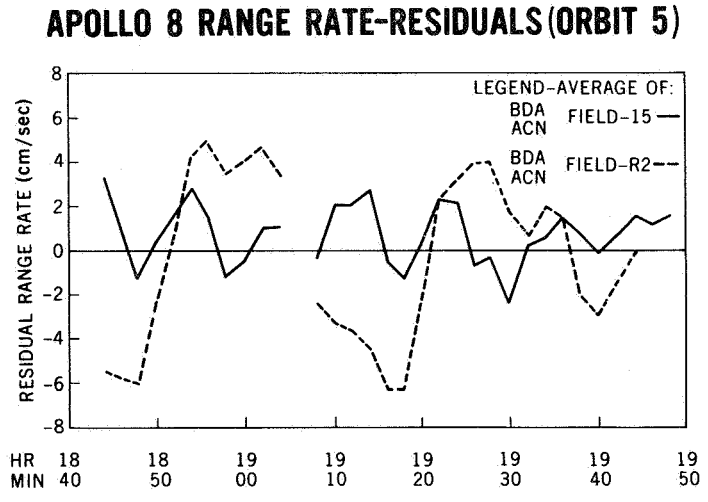


Figure 2

LOFT TEST MODEL

William D. Hibbard

CHAIRMAN:

The preceding series of papers has dealt with trajectory and orbit determinations. We shall now change the subject and talk for a while about spacecraft R&D in structures, dynamics, and control. Our first speaker is Mr. William Hibbard, who will discuss the LOFT model tests.

MR. HIBBARD:

LOFT is an acronym for a low frequency radio telescope to be used for radio astronomy observations in the frequency range of about 1 to 10 MHz.

The objective of the LOFT program is to develop the technology required for the third generation of radio astronomy satellites, assuming that RAE 1, which has been generating low-resolution radio maps of the radio sky for 17 months, is the first generation and that a high resolution interferometer will be the second. The third generation will provide the space radio astronomer with his first true telescope, that is, an instrument with good resolution, high gain, and the ability to observe a particular source continuously for an extended period of time. Figure 1 compares RAE 1 with the present concept of LOFT, a spinning paraboloidal antenna about one mile in diameter.

The enormous size is an inescapable consequence of the frequency, gain, and resolution requirements. The spinning paraboloid offers a number of significant advantages:

1. Simplicity, in that only an inert reflector is actually spin-deployed.

2. Broad frequency range, with no mechanical tuning.

3. Steady, continuous pointing, since the antenna beam is coincident with the spin axis.

4. Reorientation capability, made practical by the three-dimensional structure, which provides the necessary rigidity for attitude control torquing.

In Figure 1, the artist has taken the liberty of showing a solid LOFT reflector; actually, the reflector will be a net so tenuous as to seem invisible.

During 1969, progress on LOFT was made in several directions. To investigate the radio frequency design, and particularly the antenna feed, several scale models were built and tested on a pattern range. Considerable progress was made in identifying the dynamic modes of the 5-m-, 50-m-, and 1500-m-diameter LOFT models. Also, a preliminary study of the instrumentation requirements for a 50-m flight model was completed.

However, the major progress was the successful development and testing of a 5-m-diameter LOFT in GSFC's dynamic test chamber (DTC). Shown in Figure 2, this model is capable of stowing and deploying, in vacuum, a complete paraboloidal net. Although it does not include any radio frequency equipment, this model has successfully demonstrated the feasibility of the deployment concept through repeated tests, both at the contractor's plant and in the DTC. The 5-m LOFT model is compatible in size, weight, and general ruggedness with the Aerobee sounding rocket, which provides the subject of the 1970 LOFT presentation.

Figure 3 is a photograph taken during one of the DTC tests. The overall structure, including the paraboloidal net, the forestays, the small backstays, the mast compression column, and the central canister, is readily visible. The general parabolic shape of the net can be seen in outline.

As oftentimes happens, the development of the facilities was a task to rival the development of the LOFT itself. Net contour was measured by photogrammetry, which led to new camera and calibration techniques, since general photography and television were unsuitable because of the net construction. An improved speed control was developed for the DTC spin table to enhance its reliability and performance, and a special air jet exciter, rotating at high speed, was built to introduce dynamic disturbance into the net. These and other facility needs were satisfied to the requirements of LOFT, and to the benefit of many other DTC users.

As a postscript, in 1969 the LOFT program provided some technology to be utilized by the Earth-bound world as well: The United States Army is purchasing jeep-mounted, erectable antennas for field use, based directly upon the LOFT 5-m mast compression column.

RADIO TELESCOPE TECHNOLOGY

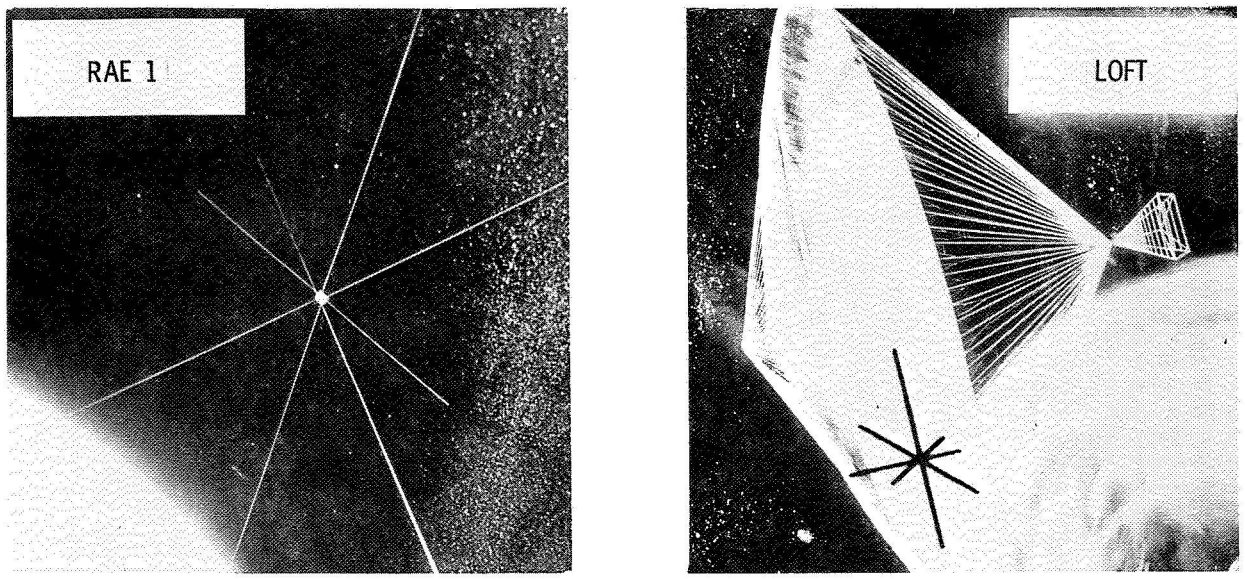


Figure 1

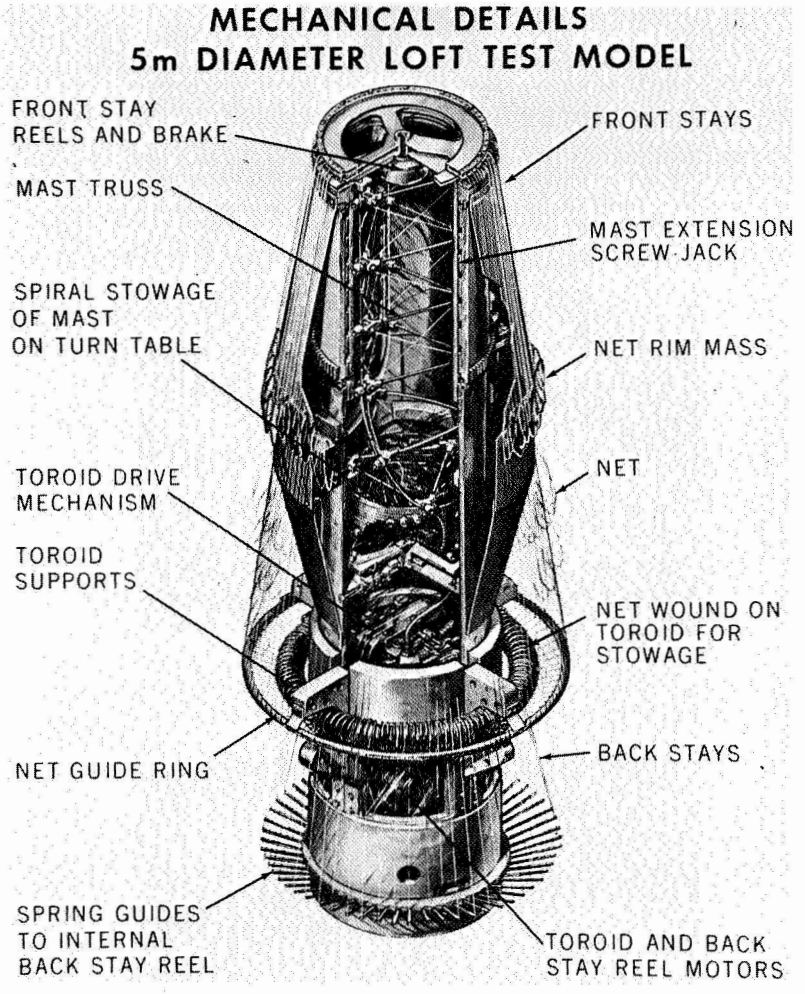

MECHANICAL DETAILS
5m DIAMETER LOFT TEST MODEL
FRONT STAY REELS AND BRAKE
MAST TRUSS
SPIRAL STOWAGE OF MAST ON TURN TABLE
TOROID DRIVE MECHANISM
TOROID SUPPORTS
NET GUIDE RING
SPRING GUIDES TO INTERNAL BACK STAY REEL
FRONT STAYS
MAST EXTENSION SCREW JACK
NET RIM MASS
NET
NET WOUND ON TOROID FOR STOWAGE
BACK STAYS
TOROID AND BACK STAY REEL MOTORS

Figure 2

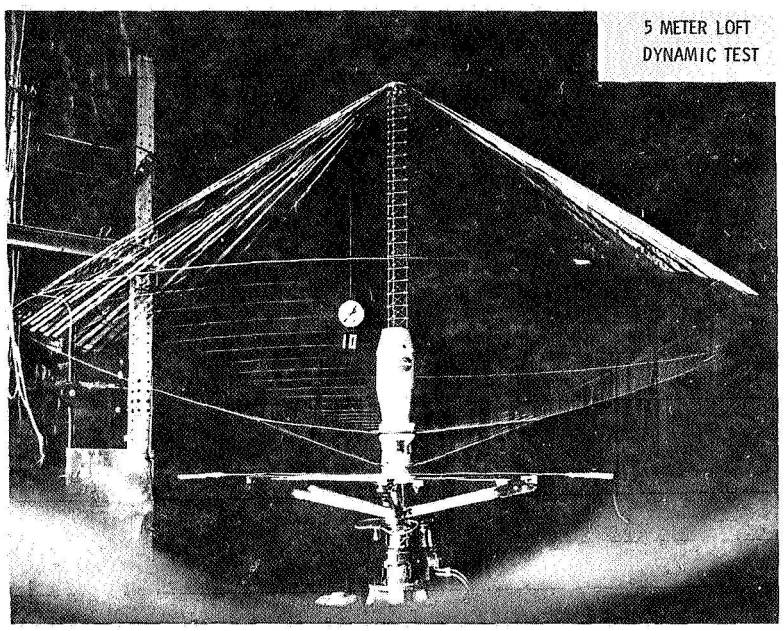


Figure 3

AN ANALYSIS OF THERMALLY INDUCED OSCILLATIONS OF STEM-TYPE BOOMS

Harold P. Frisch

CHAIRMAN:

The next speaker is Mr. Harold Frisch, who will discuss the analysis of thermally induced vibrations of STEM-type booms.

MR. FRISCH:

Before the flight of RAE 1, all satellites requiring long, extendable booms used the standard STEM-type boom shown in Figure 1. It is made from a long, thin, prestressed tape that is rolled on the drum for storage and, when extended, takes on a tubular shape.

In early 1966, I performed a static, thermal-bending analysis of STEM-type booms. Results of this analysis led me to suspect that the boom's extreme weakness in torsion could possibly lead to thermal-bending problems. As a result of this and other considerations, torsionally rigid booms were substituted in the design of both RAE 1 and ATS 5.

Before the flight of RAE 1, STEM-type booms had been used on such satellites as OGO 4 and OGO 5 for electron probes and on ATS 1, GGSE 3, 5, and 6, OV 1 through 10, and on several other Navy and Air Force satellites as gravity-gradient booms.

On each of these satellites, varying degrees of anomalous spacecraft body motion have been noted. It is now agreed that the cause of this anomalous motion is the thermally induced oscillations of the attached booms.

The basic idea behind the theory of thermally induced oscillation is quite simple. When the boom is placed in sunlight, a temperature difference will build up between the sunlit side and the dark side of the boom at a rate determined by its thermal characteristics. This naturally causes the boom to bend at a rate determined by its mass characteristics toward a position of thermal equilibrium, that is, toward the position the boom would instantaneously assume if it had zero mass.

As the boom bends, it will also twist. This causes a change in the heat flow around the boom and also a change in the position of thermal equilibrium towards which the boom is moving. It has been shown that, over a broad range of conditions, the position of thermal equilibrium can

change at a natural frequency of the system. The system is then in effect driven at resonance, and large-amplitude, thermally induced oscillations can be excited.

Recently, I have developed a computer program that has allowed me to check this theory against actual satellite data. In my analysis, the boom is assumed to be initially at rest, rigidly clamped at one end, and free to both translate and rotate at the other end. The illumination from the Sun is normal to the boom and provides the only source of disturbance.

In Figures 2, 3, and 4, some of the results of this study are shown. Each of these figures shows how nonplanar tip motion grows in amplitude as a function of time. To do this, the origin of the two-dimensional coordinate system in which nonplanar tip motion is plotted translates without rotating along the oblique time axis at a constant rate.

Figure 2 shows the predicted tip motion of the 60-ft OGO 4 boom. Note that in the absence of all external sources of disturbance, except that due to the impinging thermal energy of the Sun, the oscillations quickly grow in amplitude to about ±19 ft. This agrees well with what can be inferred from actual satellite data.

Figure 3 shows the predicted tip motion of the 30-ft OGO 5 boom. Again, the predicted ±3-ft oscillation agrees well with what can be inferred from satellite data. Figure 4 shows the effect of a change in the boom's torsional rigidity. Here, the torsional rigidity of the 60-ft OGO 4 boom has been increased by only a factor of 15. The response is stable.

Figure 5 shows three concepts for increasing boom torsional rigidity. In each concept (the Westinghouse zipper boom, the Fairchild-Hiller zipper boom, and the SPAR interlocked Bi-STEM) the boom's torsional rigidity is increased by at least a factor of 1000, a point well beyond that needed to achieve stability for all satellite boom configurations studied.

In addition to the study of past problems, this analysis has also been applied to an investigation of possible future problems. In particular, torsionally weak STEM-type booms were to be used on OGO 6, NIMBUS D, the French satellite FR B, and one of the Apollo applications experiments proposed for Apollo 19 and Apollo 20.

Each of these boom configurations was found to be highly unstable in a solar environment. They have now all been replaced by torsionally rigid booms. Of these satellites only OGO 6 has been launched. It has been in

orbit now for approximately 6 months, and, as predicted, its two 30-ft torsionally rigid booms are stable.

Furthermore, as a direct result of this analysis, it is now recommended practice to use torsionally rigid booms on all three-axis stabilized satellites that require booms for either experimentation or gravity-gradient stabilization.

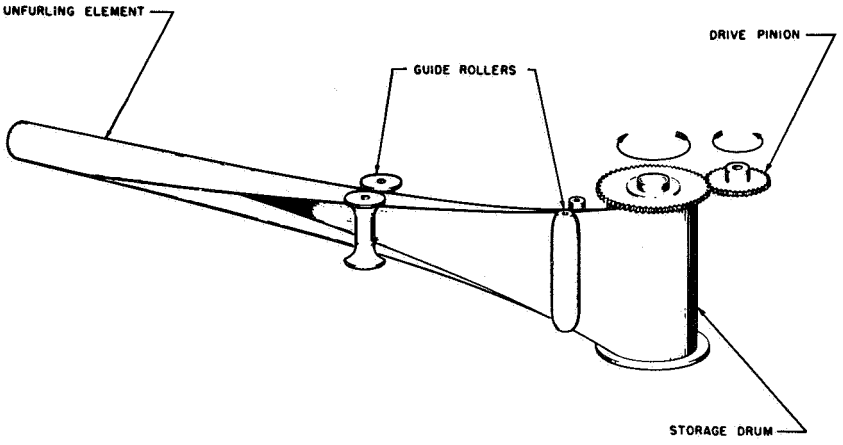
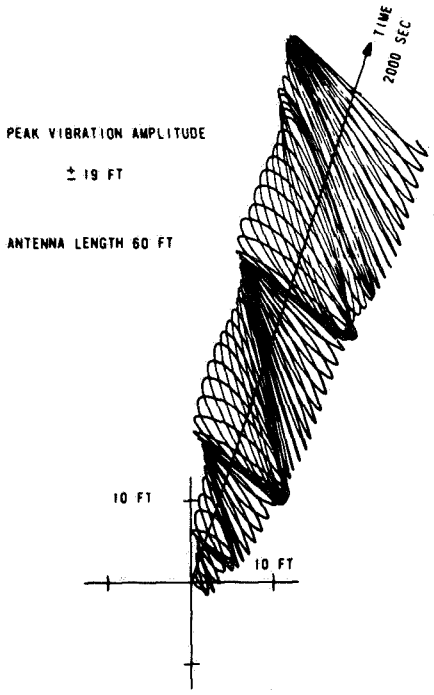
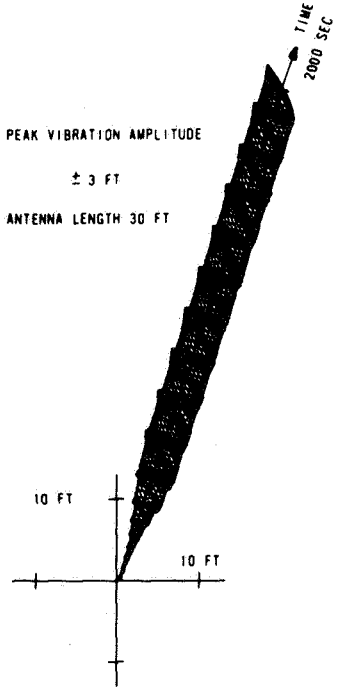


Figure 1



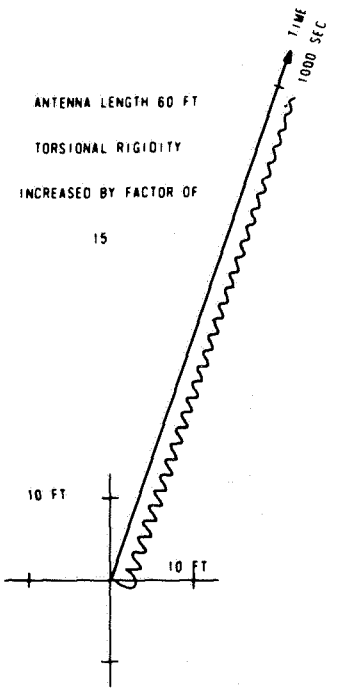
TIP MOTION OF 60-FT ANTENNA ON OGO 4 WHEN PLACED IN SUNLIGHT AT TIME ZERO

Figure 2



TIP MOTION OF 30-FT ANTENNA ON OGO 5 WHEN PLACED IN SUNLIGHT AT TIME ZERO

Figure 3



TIP MOTION OF 60-FT ANTENNA ON OGO 4 WHEN PLACED IN SUNLIGHT AT TIME ZERO

Figure 4

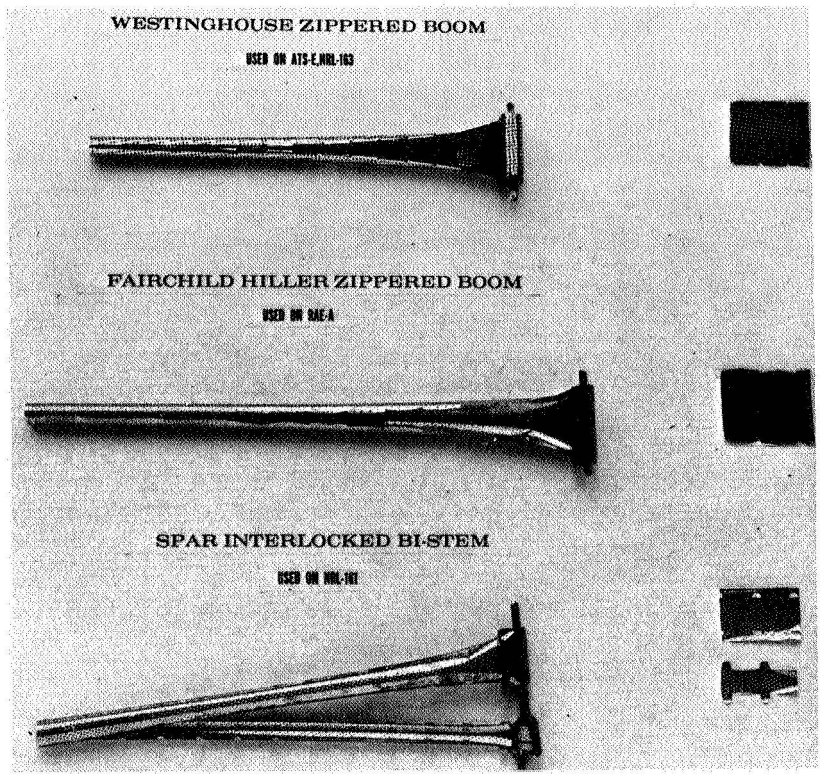


Figure 5–Torsionally rigid boom concepts presently in use.

THE NASA STRUCTURAL ANALYSIS GENERAL PURPOSE DIGITAL COMPUTER PROGRAM

Thomas G. Butler

CHAIRMAN:

The next speaker is Mr. Thomas Butler, who will discuss some applications of the NASA structural analysis general purpose digital computer program.

MR. BUTLER:

We have described the capabilities that were planned for the NASA structural analysis (NASTRAN) program. Today, I shall discuss how these capabilities are coming to fruition on spacecraft projects at GSFC, with the intention of demonstrating how NASTRAN has met its objective of being a very general analytical tool. Structural examples are taken from analyses made on the Planetary Explorer Satellite, the GSFC experiment package to be flown on the Orbiting Astronomical Observatory (OAO), and the Small Astronomy Satellite (SAS). The features of NASTRAN to be exhibited are stress analysis, thermal deformation analysis of a structure having varied materials, and vibration analysis of a multiply connected structure.

The critical factor in the design of the Planetary Explorer was weight. However, attention was first given to the distribution of the static load path and subsequently, to the sizing of individual members for strength. The satisfaction of these two stress conditions resulted in a weight reduction, a benefit of good stress design practice.

Figure 1 is a product of the NASTRAN plot routine. Only one-half of the structure is shown because it is axially symmetrical. Lines in the outer shell represent the beam supports to the solar panels, and the lines in the inner shell represent the outlines of plates. The three sections marked were the substructures under design review. Each of the three carries loads from regions remote from the axis into the stronger, stiffer inner cylinder. It is important to keep superfluous load and weight out of the outer structure because the relatively fragile solar cell panels mounted there are very sensitive to flexural strains. The dashed lines in the figure represent early design attempts, or trial configurations; the heavy solid lines depict the optimum solutions by the NASTRAN system stress analysis.

NASTRAN has considerably facilitated systems design and analysis in that the consequences of a proposed change can be rapidly evaluated.

Changes of materials, of the design of structural members, and of the arrangement of supports were all considered during this design study, and the consequences to the system weight, for example, of any change suggested to ease a local problem were easily investigated. Once the basic problem and constraints had been programmed, each trial took 20 minutes maximum on the computer. The analytical studies involved in the Planetary Explorer stress analysis took about four hours computer time.

A thermal problem associated with the GSFC experiment package to be flown on OAO A2 was a change from the ambient conditions at the factory during construction and test to the cold conditions after launch. The portion of the structure shown in Figure 2 extends from the primary mirror axially upward toward the pointing end of the optical structural assembly. Most of the structure is aluminum, but titanium is used for the upper telescope tube. The difference in thermal expansion of the two materials caused a bulging of this belt area, shortening the axial distance from the primary mirror to the secondary mirror by 0.010 in. This thermal effect was shown to be within the corrective ability of the secondary mirror focusing adjustment.

Vibration analyses are usually much more time-consuming and expensive to compute than are static analyses. The NASTRAN analysis of the SAS spacecraft reduced this problem by first setting up a refined model of the structure for static analysis in order to get a good representation of the elastic characteristics of the structure and to obtain a fine mesh model of the mass properties. An automatic internal module, to reduce the elastic and mass properties to a much coarser partition, was then used to preserve the integrity of the stiffness and mass data. The number of degrees of freedom of the model was thus reduced from 3500 to only 200. At this point, the investment in computer time was 4½ hours. Vibration analysis for the four lowest modes then took only 30 minutes more computer time. Figure 3 shows a bending mode shape and Table 1 shows the agreement of the analysis with test results. The two axial mode shapes differed somewhat from test data. Examination showed that the modeling of some of the massive equipment was concentrated too much in the centers of panels. Redistributing this mass produced good agreement with test on all counts.

In conclusion, NASTRAN is proving to be reliable, and the rate at which it is being accepted at GSFC is encouraging. Other centers are also adapting NASTRAN. One notable example being the Jet Propulsion Laboratory (JPL), where the good agreement between the NASTRAN analysis and test data on the 250-foot Goldstone tracking antenna has inspired sufficient confidence that JPL has made a NASTRAN analysis a contract requirement for the new, similar antenna scheduled for construction in Spain.

If the adoption of NASTRAN continues to grow, we shall be in a position to test how well it will meet its objective of becoming a standard method of analysis for structural problems throughout NASA.

Table 1

Comparison of Analysis and Test for SAS

Mode	Frequency (Hz)	
	Analysis	Test
1st bend	28.1	25
2nd bend	28.3	25
3rd axial	78.3	80
4th axial	95.9	90

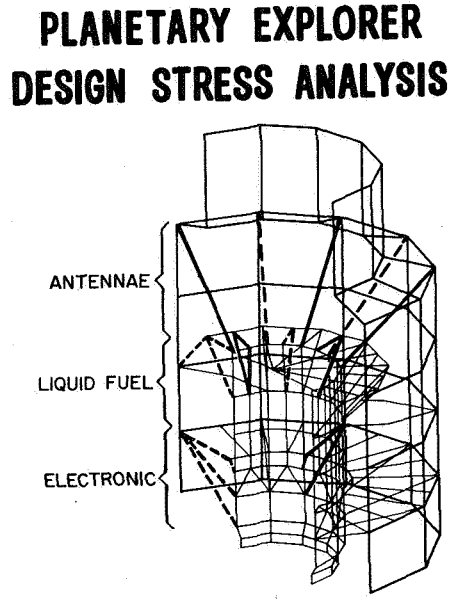
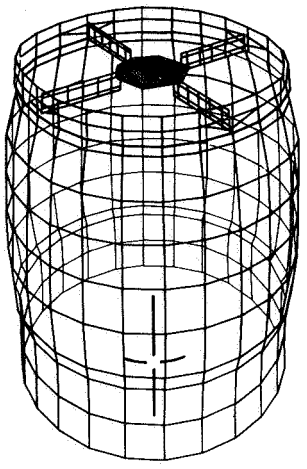


Figure 1

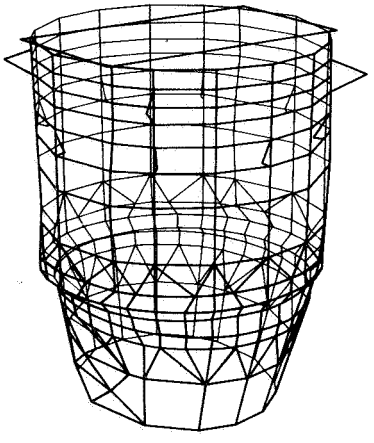
GODDARD EXPERIMENT PACKAGE THERMAL ANALYSIS



UNSYMMETRIC COLD SOAK

Figure 2

SMALL ASTRONOMY SATELLITE VIBRATION ANALYSIS



BENDING MODE

Figure 3

THE DEVELOPMENT OF A MAGNETICALLY SUSPENDED MOTOR FOR ATTITUDE CONTROL DEVICES

Philip A. Studer

CHAIRMAN:

The next speaker will be Mr. Philip Studer, who will discuss a magnetically suspended motor and its application to attitude control devices.

MR. STUDER:

The development of an electronically commutated dc motor with a magnetically suspended rotor represents a significant step forward toward the goal of unlimited-life electromechanical devices. We believe that this is the first, or one of the first, practical applications combining magnetic bearings in a motorized assembly with load-carrying capability in the tens of pounds. The absence of all physical contact between surfaces in relative motion eliminates wear and restrictions on rotational speeds. Smaller and lighter motors and momentum wheels are possible when life is not a function of rotational rate.

Figure 1 shows an engineering model in which the rotor of a "brushless" dc motor is completely suspended by magnetic forces. There are no bearings in the conventional sense. This model has a 6-lb rotor and an 8-oz-in. torque capability; it operates at speeds up to 10,000 rpm and has very low rotational losses. Electronic commutation and control of the motor make it possible to accelerate in a controlled manner in either direction of rotation.

The suspension system consists of a pair of symmetrical, cylindrical electromagnets, one at each end of the rotor and both operating in the attractive mode. The axial position of the rotor is sensed photo-optically, and the relative magnet excitation is controlled differentially to maintain equal rotor spacing between the two fixed magnets. An electromagnetic sensing coil provides a rate-of-axial-motion signal which is introduced to the servo system for stabilization.

In 1842, Earnshaw proved that complete three-axis support by passive (i.e., permanent magnet) fields was mathematically impossible. With the active axial control, however, the passive magnetic forces can provide both radial and torsional support. The rotor aligns itself parallel to and centers itself on the axial centerline in order to satisfy the condition for minimum reluctance; no control is required. The relative axial and radial load

capability is a complex function of gap geometry and magnetization characteristics of the pole-face structure. The model exhibits approximately 2-*g* radial and 4-*g* axial restraint with a stiffness exceeding 1000 lb/in. in both directions at a power expenditure of about 10 W/bearing.

The motor is a key element in this development because conventional designs impose large destabilizing forces in the motor air gap. The motor designed for this development has an ironless armature which, in addition to eliminating static forces between the rotor and stator, exhibits no hysteresis (breakaway) torques. Thus the motor does not compromise one of the outstanding advantages of the magnetic bearings, that of negligible stiction torque. The integral-motor-bearing approach to the problem was considered a necessary factor to develop a practical system; magnetic bearings without drive and rotational control capability are little more than laboratory curiosities. The low inductance of the iron-free armature provides extremely fast torque response.

This motor design, although heavier than more conventional designs of similar torque, has virtually all of its mass rotating, making the whole system particularly well suited for momentum and reaction wheel systems where high inertia is required. A single permanent magnet is sandwiched between a pair of identical multipole soft-iron pole pieces which are centrally mounted in the rotating assembly. Between the faces of these poles is a stationary disc-type armature with multiple-wound coils molded in. It is supported around its periphery by the stationary structure. All of the rotor elements are passive, including both motor and bearing rotating parts; the windings, from which heat must be removed, are stationary so that heat may be conducted away. The absence of power dissipation on the rotating assembly is very advantageous for vacuum operation where only radiation can remove heat from the rotor.

The weight penalty and the power required by a magnetic bearing system are the apparent disadvantages when compared with the 1-oz (or less) ball bearings which it replaces. The weights of both momentum wheels and motors are, in general, inversely proportional to the speed at which they can be operated. In the design of devices where speeds heretofore were restricted to insure sufficiently reliable long-term life, this restriction may be removed. Noncontacting systems have a lifetime that is independent of speed and consequently may be run at higher speeds, thus allowing overall weight savings even though the weight of the suspension elements is significant.

Sufficient load carrying capability is required to permit testing in a 1-*g* field, but the average power required to stably support the rotor may well

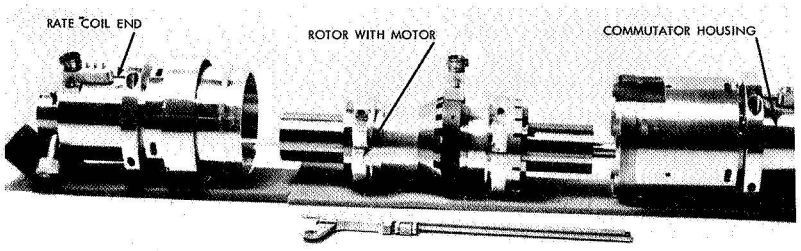
be reduced at zero-*g* conditions. The noncontacting system eliminates lubrication headaches and the need for a hermetically sealed case. (The weight of the sealed case has been the upper bound for reaction wheel sizing calculations.) Additionally, vacuum operation is more efficient than sealed-system operation, which requires pressurized atmospheres to reduce lubricant losses. In short, we believe this development presages electromechanical rotating devices having solid-state-type reliability and which will be as light or lighter than current devices. Such advances are needed to meet adequately the control-systems requirements of long-term planetary exploration and operational communications satellites and man-rated space stations.

The negligible hysteresis of both the motor and magnetic bearing results in extremely small breakaway torques. These torques are one of the ultimately limiting factors in high-accuracy pointing systems. Therefore, what may well turn out to be the most significant aspect of this development is the application to telescopes and other high-accuracy, optical communications systems.

The second model of this type is currently being improved by the addition of an inertia wheel to a prototype momentum storage device.

A great deal of work remains to be done to move from this empirically designed model to achieve the capability of confidently designing to a wide variety of application requirements. To this end, we have initiated a computerized dynamic analysis and optimization effort which, coupled with the testing and experimentation on this engineering model, will move us rapidly toward that goal.

The eventual nonspace applications of this development are those in which the reliability, the cost of maintenance, or the difficult environment are challenging. These are found in aircraft, nuclear reactors, and remote-site equipment. One of the significant features is that, although the mechanical clearances required for efficient magnetic bearings are small, they are large compared to thermal expansion effects, and the machining accuracy required for the bearing surfaces is not as demanding as that for precision ball bearings.



EXPLODED VIEW

MAGNETICALLY SUSPENDED MOTOR

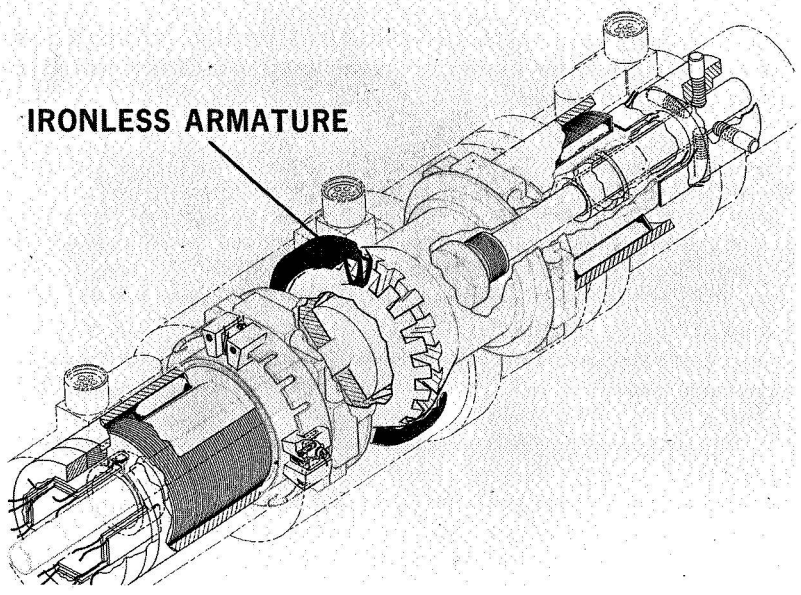


Figure 1

LOW-INPUT VOLTAGE CONVERSION FROM UNCONVENTIONAL PRIMARY (RTG's) AND SECONDARY (BATTERY) SOURCES

Edward R. Pasciutti

CHAIRMAN:

Resuming the area of spacecraft R&D, we will now cover power, thermal control, and materials. The next speaker is Mr. Edward Pasciutti, who will talk about low-input-voltage conversion.

MR. PASCIUTTI:

The primary source of power for some proposed deep space missions will be provided by radioisotope thermoelectric generators (RTG's). RTG's, however, have limited power, and storage devices may be needed to supply intermittent peak power demands. Considering electrochemical storage, the needed reliability (lifetime and cycle life) probably cannot be achieved using many series-connected cells. A single-cell approach, combined with low-input-voltage conversion (LIVC), enables precise battery charge-discharge monitoring and control as well as simplified redundant battery switching.

Figure 1 is a schematic and block diagram of the low-input-voltage converter regulator (LIVCR). The schematic portion is the voltage stepup dc-dc converter section containing new circuit developments whose performance will be discussed. The LIVC changes the low dc voltage to a higher, usable dc voltage. By applying the low dc voltage to a transformer primary and reversing the primary-side potential each half period by means of two transistor switches, a square wave, stepped-up ac voltage is generated on the secondary side. This is the push-pull inverter process. This ac output is then rectified to obtain the higher dc output voltage. However, in the voltage stepup, push-pull oscillator-inverter, the "double-on" condition created by the transistor storage time raises a major reliability problem, especially at the high current levels associated with LIVC.

Figure 2 illustrates the waveforms of a new total-cycle, load-related, base-current shaping technique which essentially eliminates the harmful effects of transistor storage time. The inverter switchover (at time t_2) occurs when stored base charges are minimal. Thus, storage time is minimized. The load-related base-current reduction of the off-going transistor and the load-related base-current overdrive of the on-coming transistor are in two distinctly separate time intervals. This minimizes the harmful simultaneous on-time conditions encountered with transistor storage time overlapping.

The overdrive of the on-coming transistor speeds circuit rise time. To implement the new crossover method required the addition of only one component, an inductance of prescribed value in the frequency-determining loop. This generates the loop current i_L, which shapes the base current through base transformer current sharing. The optional circuit switching achieved is smooth, fast, complex in theory, simple in implementation, and reliable in practice.

Another problem area is that of inverter half-cycle unbalance. This leads to magnetic saturation of the output transformer core during cycling, which could result in failure-inducing current levels, particularly when the inverter is energized by a low-impedance battery source. A means of decoupling the source from the short-circuit loading occasioned by core saturation was needed. By using a secondary winding, the primary winding magnetizing current requirements of the output transformer were separated from the load-related current-feedback base drive. With this technique it is possible to reliably use efficient current feedback in low-impedance, battery-energized, inverter-converter applications.

Figure 3 illustrates two typical output transformer saturation conditions which can occur during inverter cycling. Note particularly that in each case the base drive was terminated immediately upon occurrence of saturation, automatically decoupling the source from the short circuit loading of output transformer saturation, and that reliable, continued operation was achieved even in the presence of severe half-cycle unbalance conditions.

The circuit design developments have improved the inverter-converter reliability by several orders of magnitude and reliable 5- to 10-year lifetimes, whether energized by high- or low-impedance sources, are now feasible.

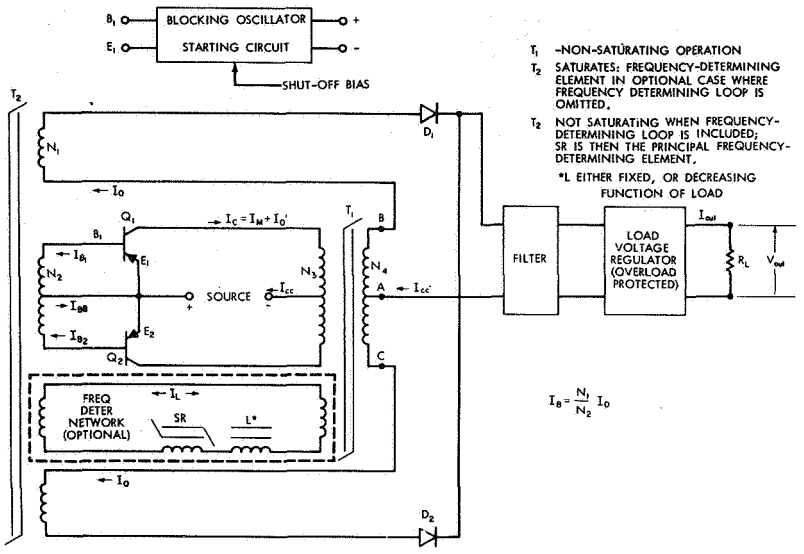


Figure 1

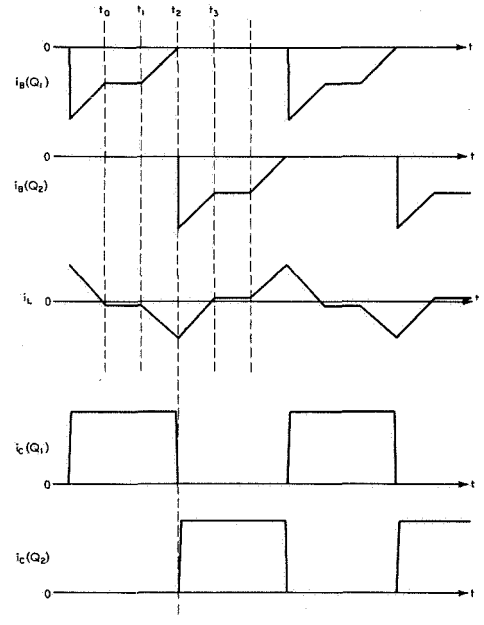
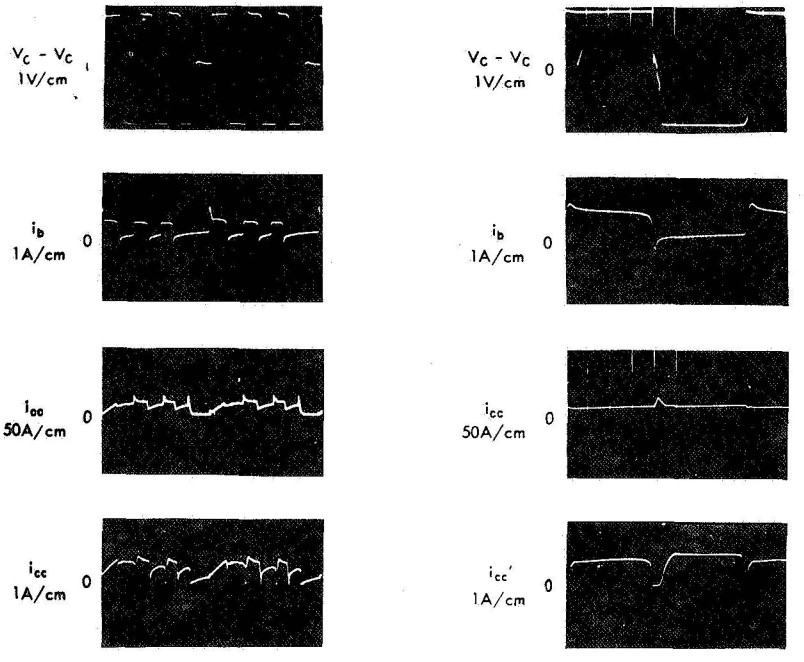


Figure 2



HALF-CYCLE SATURATION CONDITION

E_{in} = 1.5 V
I_{in} = 22.1 A (Average)
E_o = 26.146 V Regulated
I_o = 0.875 A
E_{ff} = 69%
FREQUENCY = 998 cps.
R_L = 30 Ω

END OF HALF-CYCLE SATURATION CONDITION

E_{in} = 1.5 V
I_{in} = 21.6 A (Average)
E_o = 26.21 V Regulated
I_o = 0.875 A
E_{ff} = 71%
FREQUENCY = 1168 cps.
R_L = 30 Ω

Figure 3

USE OF THE THIRD ELECTRODE AND THE RECOMBINATION ELECTRODE IN SPACECRAFT BATTERIES

Floyd E. Ford

CHAIRMAN:

Let us continue with the next paper, by Mr. Floyd Ford, which considers using the third and recombination electrodes in spacecraft batteries.

MR. FORD:

One of the most difficult problems encountered in the design of power systems using nickel-cadmium batteries is one of charge control.

Methods available to the designer in the past have included low-rate, constant-current charging; temperature-compensated, voltage charge control; and, recently, third-electrode charge control.

With constant-current charging, there has always been a severe restriction on battery capacity utilization during a given shadow period. Voltage-control charging allows for greater utilization of battery capacity, but choosing the exact voltage limit is very critical. There are the ever-present problems of thermal runaway, lot-to-lot variations in cell characteristics, changes in battery characteristics with life, and the possibility of high internal cell pressures.

The third-electrode concept overcomes the disadvantages of the voltage control and the constant-current charging. However, it has limited use in near-Earth-orbit satellites. This is due to the relatively short shadow periods that the satellite experiences during the orbits. As a result, the third-electrode voltage level used for charge control does not always decrease below the threshold necessary to allow charging when the solar array reacquires power.

A basic nickel-cadmium cell, shown in Figure 1, consists of a number of positive plates that generate oxygen gas when overcharged, a number of negative plates that serve to recombine the oxygen gas that is generated, and a third electrode that is sensitive to the oxygen pressure in the cell and which provides an electrical signal for charge control.

Through the recent addition of a recombination electrode, the subject of this presentation, we have sought to enhance the gas recombination rate within the cell and to overcome the problem identified above.

The basic operation of this cell is shown by Figure 2. When a cell is charged and overcharged, oxygen gas is generated from the positive plates, resulting in a pressure rise within the cell. This pressure rise is indicated by the third-electrode oxygen-sensitive device, which provides an electrical signal that can be used for charge control.

The third-electrode voltage signal is very reflective of the cell's internal pressure characteristic; consequently, only one profile is shown in the figure–that being the pressure in the cell. When a signal level of 300 mV is reached, the charge current can be reduced or terminated, depending on the specific application in the satellite. In most cases in satellites the charge current is not completely terminated but is reduced to a fairly low overcharge rate.

With the negative electrode alone, the recombination of oxygen is fairly slow, and a pressure decay would result, as indicated by the dotted line which is also indicative of the third-electrode voltage's being maintained above the threshold level. By the introduction of the recombination electrode, the oxygen recombination rate has been sufficiently advanced that the oxygen pressure in the cell will begin to fall off very rapidly during the discharge period, or the shadowed portion of the spacecraft orbit. As a result of the recombination electrode's causing the pressure to decay, the third-electrode signal would also follow the same type profile. It is readily apparent that the signal used for charge control would then be below the threshold level by the time the solar array acquires the Sun for the next charge period, thereby reinitiating the battery charging cycle.

The concept presented has proven to be a very reliable method of charge control. Two of the more important advantages of this concept are that the charge control is independent of the battery terminal voltage and of the battery capacity utilization.

The design of spacecraft batteries with this type of charge control concept is applicable to all types of spacecraft missions; both near-Earth orbits or synchronous-type missions can benefit markedly.

Probably the most important single achievement is that the heat generation within a nickel-cadmium battery is reduced to the lowest possible value, thereby providing a significant increase in the life capability and reliability of such batteries.

Batteries with the third electrode, or the signal-producing electrode, were first flown on the Orbiting Astronomical Observatory (OAO). These

signals were used in a passive mode for charge control. That is, they were not in an automatic loop with the battery charger. They were used by the ground crews to monitor the overcharge status of the battery.

Cells containing the recombination electrode, along with the signal electrode, are now being designed into the nickel-cadmium battery system for the Apollo Telescope Mount (ATM). The signal-electrode voltage will be used in an active mode of control–that is, in the feedback loop–and it will be the prime method of charge control for this spacecraft.

It has come to our attention during the recent months that there has been a spin-off from this effort. The National Institutes of Health is developing the concept, as presented, to be used in an implantable heart-assist device for medical applications.

In conclusion, the flexibility and advantages realized from a power system utilizing the concept presented here promises to make batteries containing the third electrode and the recombination electrode the work-horse of the space program for the next decade.

CHAIRMAN:

Are there any questions or comments for Mr. Ford?

DR. CLARK:

I assume that this is a sealed cell, at a pressure of 15 psia. I am wondering, what is the increase in pressure at the 300-mV level?

MR. FORD:

Yes sir, your assumption is correct. It is a sealed, nickel-cadmium cell, which is, of course, a hermetically sealed unit. The pressure that we would expect in this operation would be in the range of 15 to 20 psia. In normal, near-Earth orbits, using this regime, we would see a pressure drop during a 35-minute dark period of approximately 5 to 8 psia.

The cells are always under pressure, as indicated in the display we have here this morning. But what we are trying to do is to create pressure cycling in the cell such that the pressure cycle will be sufficient to cause an electrical signal cycle in conjunction.

DR. CLARK:

With the recombination electrode, what, approximately, would be the time to achieve the pressure decay?

MR. FORD:

About half of this, in other words, of the order of 17 minutes. It is not quite that clear-cut because you can tailor this recombination electrode to give you any pressure decay rate. There is another variable that was not mentioned: You can put a resistance in the leg of the recombination electrode.

So, it has a lot of flexibility that you can consider in the initial design. For instance, I may not use the same cross-sectional area of recombination electrode on a synchronous orbit that I would use on a near-Earth orbit, but the concept is applicable.

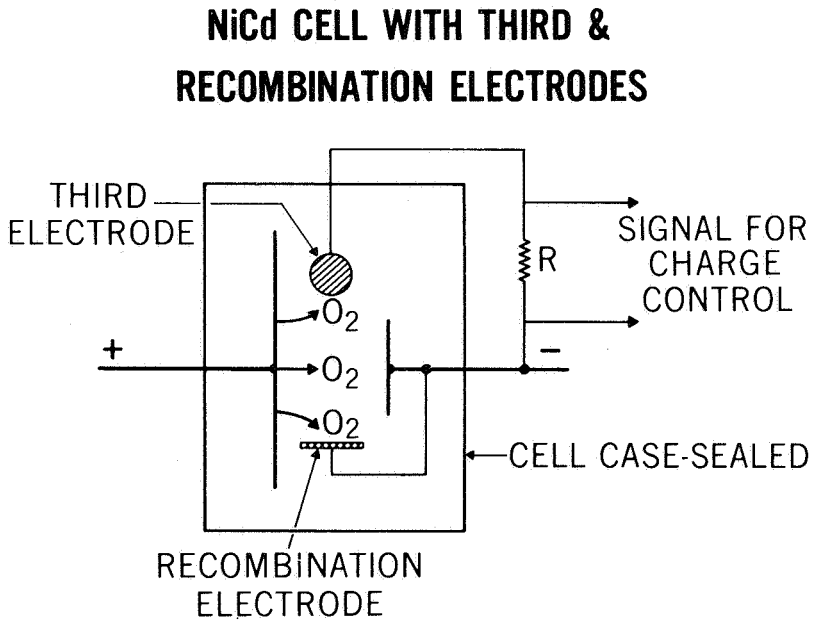


Figure 1

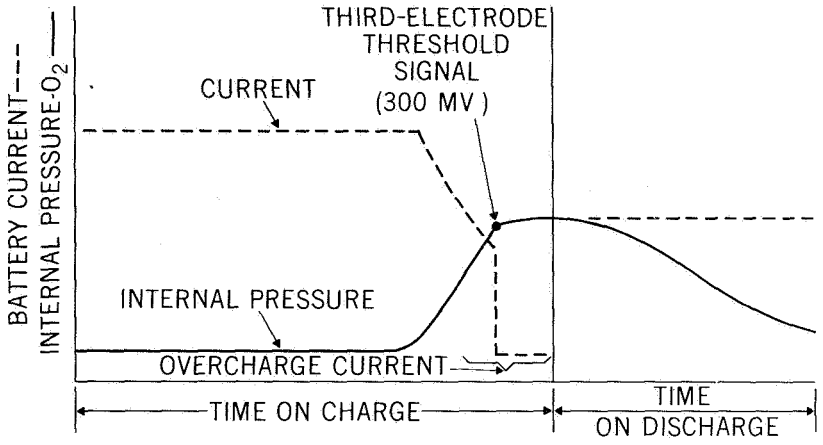

BATTERY CURRENT---
INTERNAL PRESSURE-O_2 ——
THIRD-ELECTRODE
THRESHOLD
SIGNAL
(300 MV)
CURRENT
INTERNAL PRESSURE
OVERCHARGE CURRENT
TIME ON CHARGE
TIME
ON DISCHARGE

Figure 2

RESULTS FROM THE ATS 3 REFLECTOMETER EXPERIMENT

James B. Heaney

CHAIRMAN:

The next speaker will be Mr. James Heaney, who will discuss the results of the ATS 3 reflectometer experiment.

MR. HEANEY:

In November 1967 a reflectometer experiment was launched aboard the ATS 3, which was placed in a synchronous orbit 35,780 km above the South American continent. The primary objective of the experiment was to test the relative durability of mirror surfaces in a space environment that could not be adequately simulated in the laboratory. The experiment was also designed to provide degradation rate information about Alzak, the material that comprises the entire outer shell of the OAO spacecraft. This information was to be supplied prior to the launch of the OAO A2. Most of the samples chosen were candidate coatings and substrates for use in constructing space-qualified solar concentrator mirrors. Many of them are also important for satellite temperature control applications and as protective coatings for orbiting telescopes and other highly reflective optical surfaces.

Coating experiments flown in the past have generally employed calorimetric techniques to study surface degradation rates. The test samples were thermally isolated from their surroundings so that changes in their radiative properties, solar absorptance and thermal emittance, could be deduced from temperature variations. However, the experiment described here differs from these in that it employed a reflectometer to monitor periodically, in five discrete wavelength bands, the reflectance of each test sample relative to a standard surface. This arrangement permitted study of spectral variations in the degradation process and aided in identification of the causes and mechanisms of damage.

The reflectometer experiment is shown in Figure 1. It was located on the exterior wall of the rotating satellite in a position that permitted the exposure of each sample. All of the test samples were highly reflective specular materials having aluminum or silver surfaces either uncoated or coated with silicon oxide, silicon dioxide, aluminum oxide, or magnesium fluoride. The materials to be measured were positioned in a fixed circular array containing 18 test surfaces, a high-reflectance standard, and a low-reflectance standard. Relative specular reflectance measurements were made periodically by scanning over the test samples with a rotating arm and comparing the reflectometer outputs for each sample with those obtained from the stable high- and low-reflectance standards.

Eight of the test surfaces were protected by fused silica shields from essentially all ultraviolet irradiation shorter than 160 nm, all electrons and protons, and all dust particles. Each of these protected surfaces had a mate which was unshielded and exposed to the total environment. The presence of these fused silica shields, combined with the capability for spectral resolution, enhanced our understanding of the degradation problem by better defining the causes of damage.

Because of the large number of samples flown, it is not feasible to include all the results in this brief presentation, but the data shown in Figure 2 for Alzak and aluminum oxide over aluminum are typical. In general, for the dielectric-overcoated samples the shielded samples degraded almost as much as did the unshielded ones. The silicon dioxide- and aluminum oxide-coated aluminum were the most stable of this type. In each case, the rate of loss of relative specular reflectance diminished as time progressed and was essentially limited to wavelengths less than 1200 nm.

The Alzak data confirmed laboratory tests that showed this commercially available material to be potentially unstable. The solar absorptance of the unshielded Alzak sample increased from 15% initially to 26% after 2000 hours of solar exposure in orbit. Combining the flight data with laboratory results enabled the development of an acceptance criterion for selecting stable Alzak to be used in fabricating the OAO skins. Failure to meet this criterion resulted in the removal of two unacceptable skins from the OAO A2 spacecraft. This criterion is presently being applied to the latest OAO.

The uncoated and unshielded aluminum samples flown aboard the ATS 3 reflectometer exhibited a large loss of specular reflectance in the 300- to 400-nm-wavelength region. The rate of this reflectance loss increased rather than decreased with time. On the other hand, the shielded samples changed very little, indicating that the bare metal surfaces were not damaged by ultraviolet irradiation. The unshielded aluminum samples exhibited negligible losses of relative specular reflectance for wavelengths longer than 1200 nm.

The results demonstrated that vapor-deposited aluminum, coated with silicon dioxide or aluminum oxide, would be a suitable surface for a solar energy collector mirror. The use of the fused silica shields has made it possible to conclude that ultraviolet irradiation is primarily responsible for the damage observed in the dielectric-overcoated samples, although a noticeable amount of damage is caused by electrons, protons, and other environmental factors acting in combination. The results also minimized the importance of dust particles as a source of damage, although large numbers of micrometeoroids, smaller than any yet detected, may explain the unusual

degradation behavior exhibited by the unshielded aluminum samples, since surface roughening would result in a wavelength-dependent loss of reflectance to this specular reflectometer.

Converting this reflectance data to solar absorptance values showed that coating aluminum with relatively thick but negligibly absorptive dielectric films can produce thermal control surfaces having α_s/ϵ ratios comparable to that of white paint while being more stable than previously tested white paints.

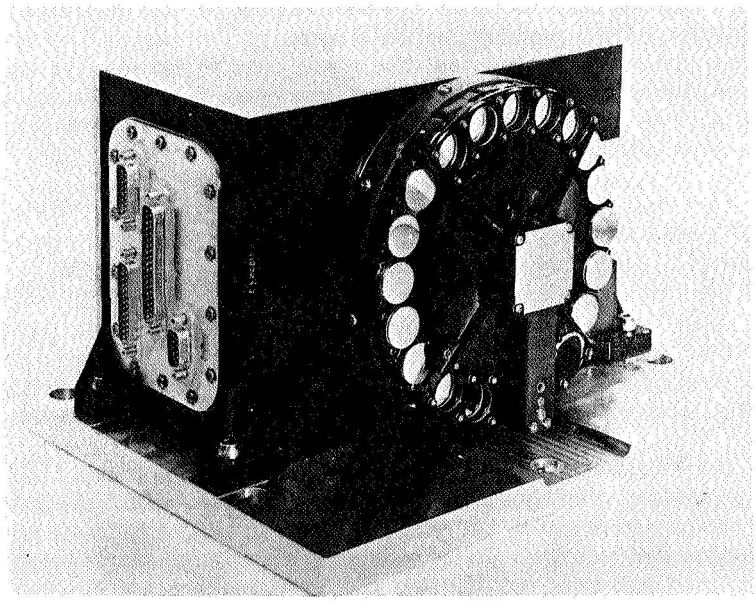


Figure 1–The ATS 3 reflectometer experiment.

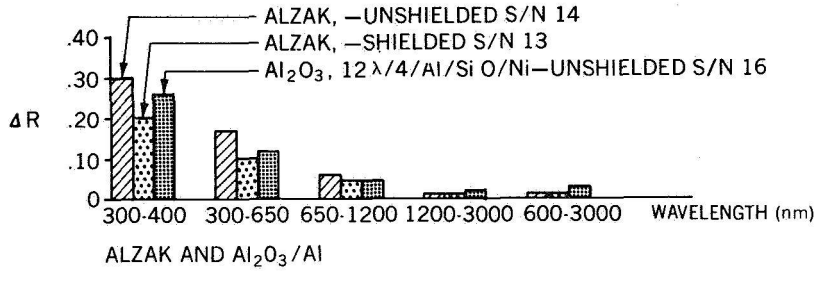


Figure 2

THE EVALUATION OF A PLANAR RADIOISOTOPE THERMOELECTRIC GENERATOR THERMAL MODEL

Joseph F. Fry

CHAIRMAN:

The next speaker will be Mr. Joseph Fry, who will present information on the evaluation of a planar radioisotope thermoelectric generator thermal model.

MR. FRY:

The planar radioisotope thermoelectric generator (RTG) model is a new space concept. For the first time, the full-scale model incorporates a planar, unidirectional radiator that permits a flexible spacecraft configuration. The planar RTG can be mounted either on the surface of a spacecraft or in a deployed position. It utilizes, for the first time, a spherical fuel capsule configuration and multifoil insulation around the heat source. The radiator is unique in that it is fabricated from a large block of hot-pressed beryllium. The model also has high-temperature electrical heaters and a thermopile consisting of 12 silicon-germanium thermoelectric couples and 100 thermal analog couples arranged about the center of the radiator.

Figure 1 is a schematic diagram of the planar RTG thermal model. The upper portion shows the large, square, beryllium radiator, which is 32 in. on a side. The lower portion shows the locations of the thermopile, the heat source, and the multifoil insulation that surrounds the heat source.

The objectives of our program were to demonstrate the feasibility of the planar RTG as a power source, to verify the analytical predictions of the temperature profiles across the radiator and across the multifoil insulation, and to verify the thermoelectric performance of the 12 silicon-germanium thermoelectric couples. The performance of the planar RTG thermal model is shown in Table 1.

Table 1
Thermal Model Results

Characteristic	BOL* Predicted	BOL* Actual	63 hr	302 hr
Power output (W)	9.5	9.65	8.5	8.2
Efficiency (%)	5.2	5.5	5.0	4.9
ΔT across multifoil insulation (°C)	690	650	650	650

*Beginning of life.

The beginning-of-life (BOL) power output was predicted to be 9.5 W at an efficiency of 5.2%. The temperature difference ΔT across the multifoil insulation was predicted to be 690° C. In order to verify these objectives, the model was operated in a 12- by 15-ft vacuum chamber, located at the Test and Evaluation facility at GSFC. The average hot-junction temperature was 805° C, and the average cold-junction temperature was 300° C. At BOL, the power output was 9.65 W, the efficiency was 5.5%, and ΔT across the multifoil insulation was 650° C. After 63 hours of operation, the power output was 8.5 W at an efficiency of 5%. After this initial burn-in period, the power output and efficiency varied slightly during the remainder of the test. At 302 hours, the power output was down to 8.2 W at an efficiency of 4.9%.

The thermal model generator met the predicted design goals and performed satisfactorily for 302 hours. The BOL goal of 9.5-W power output was met with a value of 9.65 W. The temperature profiles across the radiator and across the multifoil insulation were within 7% of the predicted design values. The test was terminated after 302 hours, because the goals were met and because of scheduling problems for the vacuum facility.

In the future, we will continue to use the thermal model as a test bed for new high-temperature silicon-germanium module concepts and previously developed segmented concepts.

DR. CLARK:

What are the major causes for the decrease in efficiency during the vacuum test?

MR. FRY:

One of the major causes is the increase in internal resistance, and it is one of the quantities that we monitored. The module's internal resistance went up from 58 mΩ at BOL to approximately 70 mΩ at the end of the test. Internal resistance and the burning-in of the electrodes at the hot and cold junctions are characteristics of the material used.

DR. CLARK:

What are the prospects for reducing this rate of decrease?

MR. FRY:

I think the prospects are good because we have various research programs in progress that are working on the problem of the contact resistance between the hot electrodes and the thermoelectric material. Therefore we

are progressing toward decreasing the contact resistance and improving the materials used.

To date, we have noticed in tests of many of the generators and modules that, during the first 50 to 100 hours, there is a burn-in period during which the efficiency and the power output drop slightly and then tend to level off. There will always be a certain amount of degradation during the burn-in period.

MEMBER OF THE AUDIENCE:

How would the radiator requirements compare with those for a solar array producing an equivalent amount of power?

MR. FRY:

This particular generator, as I have pointed out, did not have a full complement of silicon-germanium thermoelectric couples; it had 100 analog thermocouples. With a full complement of thermocouples, this 32-in.-square radiator would produce 90 W, and this compares favorably with the 9.5 W (BOL) that we have gotten from the 12 couples. I did not make a direct comparison between the 90-W generator using a 32-in.-square radiator and a comparable solar array. Therefore, I do not have the dimensions of the solar array.

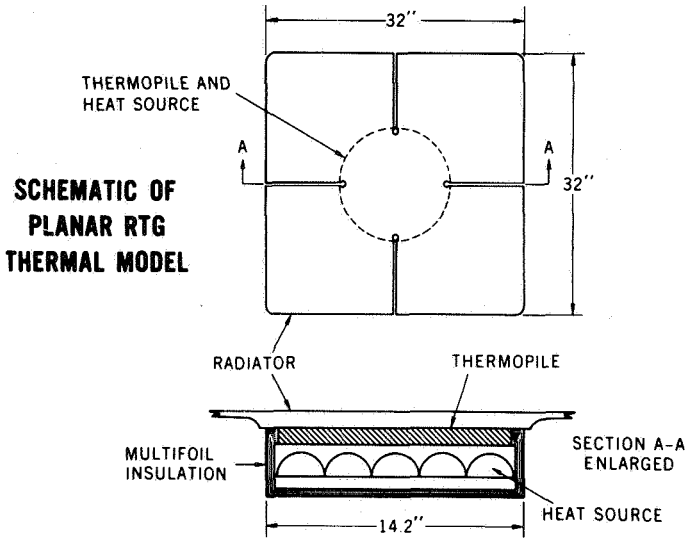


Figure 1

A SPACE APPLICATION OF A CRYOGENIC HEAT PIPE

Robert A. Callens

CHAIRMAN:

Our next speaker is Mr. Robert Callens, who will discuss a space application for a cryogenic heat pipe.

MR. CALLENS:

High-temperature and ambient-temperature heat pipes were developed to provide highly efficient, durable, and safe heat transfer devices for space application. They can transmit nearly 500 times more heat than a solid copper rod of the same diameter and length, they operate nearly isothermally, they are self-pumping, they can be used as heat flux transformers, and they can be fabricated to varying geometries to provide optimum packaging. A number of ambient-temperature and high-temperature heat pipes have been used for space applications in the last few years. By comparison, the cryogenic-temperature heat pipe has not been studied extensively and has not been put into practice. No analytical design or performance procedures for cryogenic heat pipes are available in the open literature, although the feasibility of such pipes has been demonstrated by Haskin's experiments.

Figure 1 is a schematic of a simple heat pipe which consists of a closed tube, a capillary structure saturated with a liquid, and a vapor space. The heat is transported as latent energy in a two-phase circulation. The fluid is evaporated by heat addition in the evaporator section of the heat pipe; the vapor then flows down the vapor space toward the condenser section where the heat is removed. The vapor condenses and is pumped back toward the evaporator section by the capillary forces in the capillary structure (wick). Note that for a simple heat pipe there is no preferred direction of flow. Either end could be the evaporator or the condenser.

The integrated cryogenic isotope cooling engine (ICICLE) system, under development at GSFC, will use a regenerative, heat-driven refrigerator to produce cryogenic refrigeration for long-term space applications. Cryogenic heat pipes are proposed as the thermal link between the central refrigerator and remotely located, cryogenically cooled sensor platforms.

GSFC, with the aid of a grant to Catholic University, has undertaken production of a generalized mathematical model of cryogenic heat pipes in order to better understand their operations and limitations. Figure 2

illustrates an application of the mathematical model to predict cryogenic heat pipe performance and compare it with a copper rod of the same dimensions. The model, from which Figure 2 was plotted, considered an LN_2 heat pipe 1 cm in diameter and 30 cm long. The pipe's condenser section was programmed to be 3 cm long and its evaporator section to be 27 cm long with a zero-length adiabatic transport section. Along the curve's abscissa is plotted the heat pipe condenser temperature (T_{COND}), and along the ordinate is plotted the temperature rise from the pipe's condenser-to-evaporator sections ($T_{EVAP} - T_{COND}$).

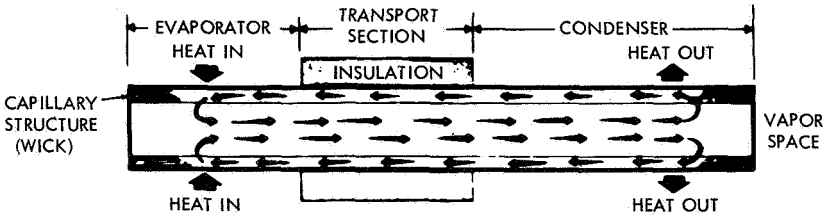
The figure contains lines of constant heat flow and lines of heat pipe performance with various wick thicknesses. The entire curve is bounded by a curve showing the performance of a solid copper rod of the same dimensions. Assume that a cryogenic heat pipe, cooled to 80° K at its condenser, transmitted 2 W of refrigeration to an IR detector connected to its evaporator. Figure 2 shows that the temperature rise from condenser to evaporator would be about 2° K; thus the IR detector would be cooled to about 82° K. If a solid copper rod were to be used instead of a cryogenic heat pipe, then Figure 2 indicates that the IR detector would be cooled to about 88° K. Similarly, for 3 W of refrigeration for a condenser at 80° K, a heat pipe would cool the detector to about 85° K, whereas a copper rod would cool it to about 93° K.

We have tested GSFC's cryogenic-heat-pipe mathematical model against Haskin's experimental results. Haskin's heat pipe was a crude device about 1 ft long that used a rayon wick and LN_2 as the cryogen. It is the only cryogenic heat pipe data available in the literature. Figure 3 shows the results of that computation. The figure shows excellent agreement between the two, especially in the area of less than 20 W, which is the area of interest for ICICLE.

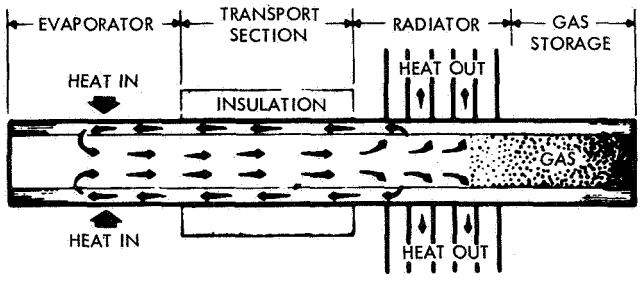
ICICLE is a potential space application for cryogenic heat pipes. The ICICLE system will be a long-term spacecraft system for cooling multiple spacecraft experiments, sensors, and so forth to cryogenic temperatures. A miniature Vuilleumier (VM) regenerative, heat-driven engine will be the central cryogenic refrigerator. The VM was patented in the United States in 1918 by Rudolph Vuilleumier, and similar thermodynamic concepts were patented by Bush in 1938 and Taconis in 1951. The ideal VM engine produces cryogenic refrigeration by He gas compression and expansion brought about by a high-temperature heat source and an ambient-temperature heat sink. The motion of two piston displacers (one hot, the other cold) in conjunction with thermal regenerators is required to produce refrigeration. Ideally, the engine has perfect mechanisms, perfect heat transfer, and no

gas friction and thus operates without mechanical energy input. Actually, an electric motor will be used to properly phase and drive the piston displacers.

Figure 4 is a block diagram of an ICICLE system. The VM engine is shown centrally located in the spacecraft with a radioisotope thermal power supply remotely located on the spacecraft's exterior. Thermal power will be transmitted from the power supply to the engine's hot cylinder via a liquid-metal heat pipe. The engine will be thermally controlled at ambient temperatures by a heat sink composed of an ammonia heat pipe and a space radiator. Highly conductive, LN_2 cryogenic heat pipes will transmit the refrigeration from the engine's cold cylinder to the sensor-platform assemblies. A thermal switch will shunt the isotope's thermal power away from or to the engine in order to turn it off or on.



SCHEMATIC OF SIMPLE HEAT PIPE SYSTEM



THERMAL CONTROL HEAT PIPE

Figure 1

CRYO HEAT PIPE PERFORMANCE

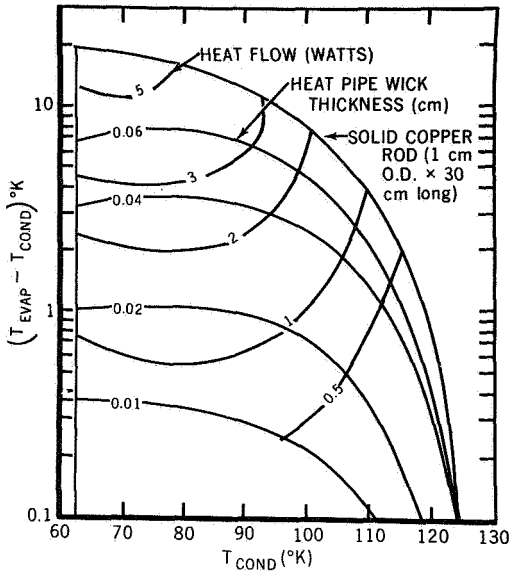


Figure 2

COMPARISON OF MODEL WITH HASKIN'S EXPERIMENT

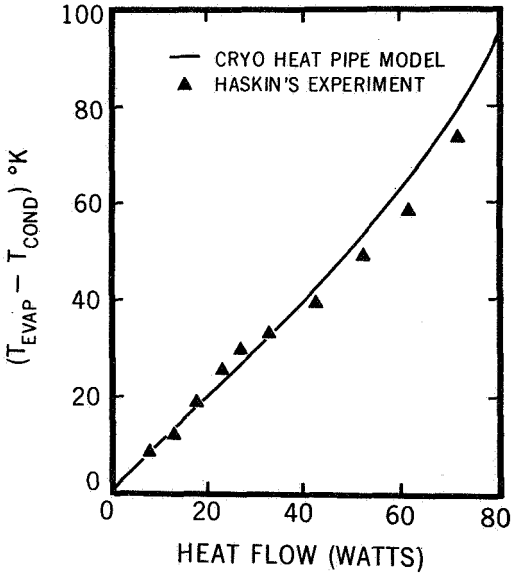


Figure 3

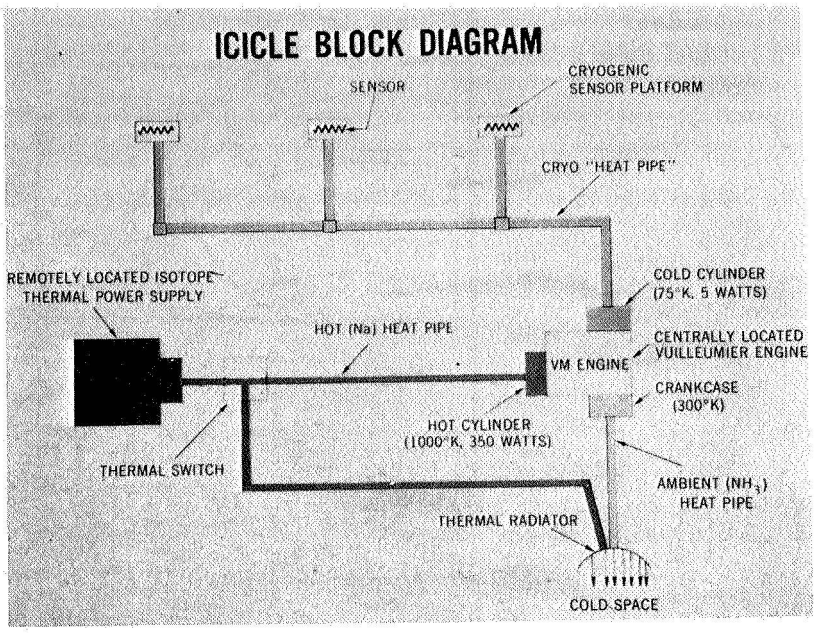

ICICLE BLOCK DIAGRAM
SENSOR
CRYOGENIC
SENSOR PLATFORM
CRYO "HEAT PIPE"
REMOTELY LOCATED ISOTOPE
THERMAL POWER SUPPLY
COLD CYLINDER
(75°K, 5 WATTS)
HOT (Na) HEAT PIPE
CENTRALLY LOCATED
VUILLEUMIER ENGINE
VM ENGINE
CRANKCASE
(300°K)
HOT CYLINDER
(1000°K, 350 WATTS)
THERMAL SWITCH
AMBIENT (NH_3)
HEAT PIPE
THERMAL RADIATOR
COLD SPACE

Figure 4

A RESIDUAL GAS ANALYSIS TECHNIQUE FOR STUDIES OF PHOTOCHEMICAL SURFACE REACTIONS OF POWDERS

Joe A. Colony

CHAIRMAN:

The next speaker will be Mr. Joe Colony, who will discuss a residual gas analysis technique for studies of photochemical surface reactions of $Al_2 O_3$ powder.

MR. COLONY:

Ultraviolet degradation of white thermal control coatings presents a considerable problem in the thermal design of spacecraft, since changes in absorption sufficient to more than double the energy input are commonly observed. A great many studies have been made in recent years to determine the mechanisms of degradation of these pigmented materials, but very little scientific enlightenment has resulted from this work. Traditional degradation theory based on color-center formation due to lattice defects, such as vacancies, dislocations, and impurities, appears inadequate to explain the empirical results with powders. Besides, the degradation of powders is often very different from the degradation of single crystals of the same material. For example, sapphire single crystals are very stable, whereas aluminum oxide powders of equal purity are very unstable. The difference, apparently, must be attributed to the most disruptive lattice defect–the solid surface and its adsorbed materials.

The available methods for studying the surface photochemistry of powders are limited, since many of the electronic measurements used for work with single crystals are not applicable to powders. Studies of a single-crystal surface cannot lead to valid conclusions for a powder where a statistical occurrence of many types of surfaces is present. Also, it happens that the properties of a surface are altered significantly when the extent of the surface is very restricted. Thus, we must study the powdered material actually used in coating preparation.

The most commonly used research method for this type of study involves surface doping with various agents and consequent checking of the effect of this on the optical stability. However, since only meager reliable results have been obtained by this method, additional means of obtaining information are needed. The residual gas analysis technique described below can be valuable whenever the adsorbed gases on the powder surface play a role in the photodegradation process.

With the residual gas analyzer the partial pressures of the low-mass species in a vacuum system can be continuously monitored. The absolute values of these partial pressures at any time are determined by a complex set of variables such as pumping rate for the species, instrument sensitivity for the species, system geometry, leak rate, sample size and surface area, and contamination. However, if the partial pressure of a component is monitored continuously before and at the onset of irradiation, the change in partial pressure will be determined solely by the interaction of the radiation with the powder surface. This observation is the basis for this study.

In order to clarify the meaning of the data which can be obtained, consider the curves shown in Figures 1 and 2. These curves, representing several general types of behavior, show the partial pressure of the gas being studied as a function of time with zero time representing the opening of a shutter. The first three curves indicate the most commonly observed behaviors including increased outgassing due to increased temperature and to photodesorption. Curves 4 through 6 (Figure 2) show three of the more more interesting and informative results with the processes responsible indicated.

In the particular case of high-purity alpha aluminum oxide, it has been observed that water, nitrogen, and carbon dioxide all react in a reasonable "well-behaved" manner. That is, they respond to give curves of the form indicating photodesorption and desorption increasing with temperature. However, the curve for oxygen has a sharp decrease at the onset of irradiation indicating the process of photoadsorption. The magnitude of this effect is sufficient to reduce the partial pressure of oxygen in the vacuum system by up to 40%. This result is independent of oxygen partial pressure over at least a tenfold range (10^{-7} to 10^{-8} mm Hg).

The significance of this phenomenon is considerable. At the very least, it indicates the presence of a very strong photoreduction potential at the aluminum oxide surface, suggesting in turn a very high sensitivity to adsorbed reducible materials. This conclusion gives direction to further surface doping studies for aluminum oxide which should result in stabilization of these coatings. Further experiments involving the effects of these pretreatments and the effects of temperature and wavelength changes on the residual gas behavior should lead to a mechanism for the instability of aluminum oxide powders.

Thus, the use of the residual gas analysis technique has been shown to be useful in studies involving photoreactions of powder surfaces and has suggested whole new areas of potentially useful research.

CHAIRMAN:

Are there any questions for Mr. Colony?

MEMBER OF THE AUDIENCE:

Would you predict the properties of beryllium oxide to be different?

MR. COLONY:

I predict it would be quite different, but it still would be very likely that, in the powdered form, it would be sharply different from those in the single crystal form. However, I could not predict what it would be.

TYPES OF PARTIAL PRESSURE CHANGES AT THE ONSET OF UV IRRADIATION

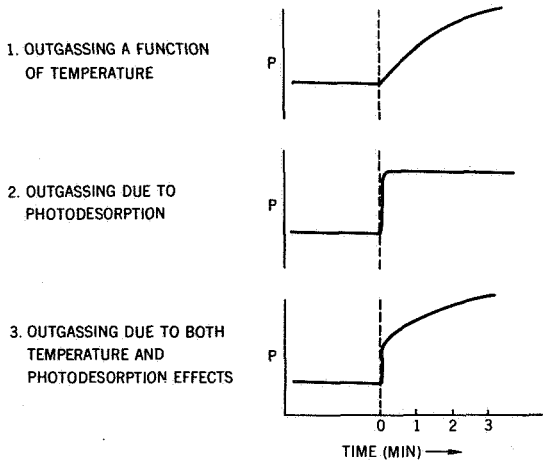


Figure 1

TYPES OF PARTIAL PRESSURE CHANGES AT THE ONSET OF UV IRRADIATION

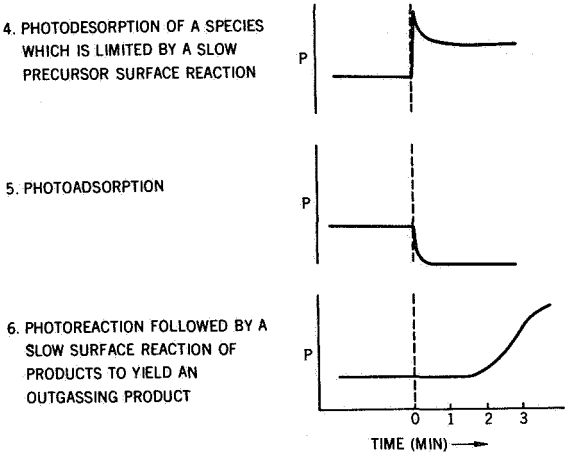


Figure 2

EFFECTS OF SURFACE FINISHING ON THE STRENGTH OF CERAMIC MATERIALS

Thomas M. Heslin

CHAIRMAN:

The next speaker will be Mr. Thomas Heslin, who will discuss the effects of surface finishing on the strength of ceramic materials.

MR. HESLIN:

Ceramics have a great deal to offer for specialized space applications. They have a high strength-to-weight ratio, they have superior strength at high temperatures under oxidizing and corrosive conditions, they are chemically stable, they do not out gas, they are radiation resistant, and they have relatively low coefficients of thermal expansion. These properties make ceramic materials complementary to, or in many cases substitutes for, conventional structural materials.

However, ceramic materials have not been fully exploited, largely because of two problems. First, relative to compressive strength, ceramic materials are comparatively weak in tensile strength. Second, they suffer from an apparent lack of predictability. Anyone who has scratched and cracked a piece of glass can tell you that the condition of the surface affects its strength. A number of ceramists feel that this surface effect also exists in other ceramics, such as the polycrystalline oxide ceramics.

The objectives of this research, therefore, have been to investigate the relationship between surface conditions and strength and the distribution of strength values and to try to develop methods for improving ceramic materials in these areas.

To date, we have worked exclusively with Al_2O_3 since it is representative of the polycrystalline oxide ceramics, and also because it is the most widely used structural ceramic for special applications.

The research program is divided into three basic areas: characterization, analysis, and development.

Characterization attempts to identify the various types of flaws that are present in a material and to establish a parameter or group of parameters that significantly characterize the material relative to strength. To accomplish this, we have used profilometer measurements, microscopes, and both scanning and transmission electron microscopes.

In analysis, we attempt to correlate strength- and flaw-magnitudes by utilizing various statistical and empirical relations. We have found that the strength data correlate reasonably well with a linear, series model–the "weakest link" model. This model states that the strength of a ceramic material is governed by the magnitude of the largest flaw present in its surface. In other words, if the material has flaws, the theory states that the largest flaws are the most significant. The object of this study is the correlation of the largest flaw populations or distributions with observed strength distributions. Despite the simplicity of the concept, the actual application of the theory involves considerable effort in flaw measuring and counting procedures.

In the area of development, we have used two approaches: flaw elimination and flaw compensation. In flaw elimination, we have used various grinding, etching, and polishing techniques such as erosion with ion beams and flash surface melting. A few of the techniques that have been tried in the second approach have involved application of compressive glazes, formation of low-expansion Al_2O_3-Cr_2O_3 solid-solution surface layers, and the bathing of glazed samples in molten KNO_3 to effect a K^+ for Na^+ exchange, thereby attempting to "stuff" the surface. An indication of the progress that has been made is given in Table 1.

Taken as a group, samples whose tension surfaces undergo a grinding and polishing procedure show slightly increased strength over untreated controls. This strength increase amounts to 5% to 10%. Ion-bombardment, surface-stuffing, and impregnation result in a decrease in strength, although some investigators have obtained increases in strength using these techniques. Glaze application brings about strength increases upward of 60%.

In conclusion we found that compressive glazing is the most effective strength-increasing technique of the ones that we have employed. Also, although our statistical studies show that strength correlates reasonably well with the "weakest link" model, more work is needed in the areas of determining what type of flaw nucleates fractures in materials such as these and what measurements significantly characterize the stress-concentrating ability of this flaw. Work is also needed in developing a rapid measuring device to count and measure the magnitude of fracture-nucleating flaws so that a parent flaw population can be properly determined.

MEMBER OF THE AUDIENCE:

How large were the samples that were measured?

Table 1

Effects of Flaw Elimination and Compensation on the Strength of Al_2O_3

Description	Strength (Flexure: 4-Point Loading)		
	Mean (psi)	Spread (psi)	Number of Samples
Flaw elimination			
Controls (as received)	30,692	5,391	8
Ground with 180-grit diamond spindle	34,055	3,690	8
Ground with spindle and finished with 30μ diamond	32,780	8,048	8
Ground with spindle and finished with 0.25μ diamond	33,171	2,943	8
Flaw compensation			
Controls (as received)	37,785	14,091	10
Covered with compressive glaze	60,280	16,057	5
Covered with compressive glaze and bathed in molten KNO_3	47,624	6,637	5
Fired in Cr_2O_3 – CrF_3	25,624	13,556	5

MR. HESLIN:

The samples were about 6 in. by 1/2 in. by 1/8 in., and the gauge area was 1-1/3 in. by 1/2 in.

MEMBER OF THE AUDIENCE:

Your compressive glaze gave you the best flaw compensation, but by the same token I noticed your spread was greater. Is there any reason for this?

MR. HESLIN:

Perhaps the glass is more sensitive to small fissures, or notching effects, than is the polycrystalline material. That is one explanation, but we have not found a firm explanation for it.

ADVANCED LOW-NOISE PARAMPS

Pio H. Dalle Mura

CHAIRMAN:

We now move to the area of spacecraft electronics. The first speaker is Mr. Pio Dalle Mura, who will speak on advanced, low-noise parametric amplifiers.

MR. DALLE MURA:

In recent years there has been increased emphasis on the use of miniaturization in the field of microwaves, especially with regard to microwave integrated circuits employing thin film microstrip circuits. The tracking and data relay satellite system (TDRSS) spacecraft has a requirement for a small, lightweight, low-noise, long-life preamplifier which dictates the use of such microwave techniques. Figure 1 shows a completely integrated, all-solid-state prototype parametric amplifier (paramp) developed for use in the TDRSS spacecraft. The paramp operates in a broadband fixed-tuned mode covering the 2.2- to 2.3-GHz frequency band. The basic paramp structure, including the ferrite circulator and the varactor circuitry, is fabricated on thin-film microstrip consisting of a 20-mil alumina substrate with a copper ground plane on one side and an etched copper conductor pattern on the other side. The solid state pump circuit is formed from conventional K_u-band waveguide components and is attached to one edge of the aluminum frame that surrounds the microstrip board.

Figure 1 also shows the complete paramp module (excluding dc power supply) with the top cover shield removed. Clearly visible on the microstrip layout are the ferrite circulator, the varactor mounted into the substrate, and the various paramp resonator circuits. The unit has an input RF connector, an output RF connector, and a dc power connector. Also shown is the circulator magnet mounted just underneath the circulator disc.

The waveguide structure attached to the upper edge of the aluminum frame is the solid-state pump source consisting of a 14-GHz Gunn Effect oscillator employing a commercially available gallium arsenide Gunn Effect device. This type of Gunn Effect device has been operated continuously by the manufacturer for well over one year with no signs of performance degradation. After earlier attempts to include the pump oscillator on the microstrip substrate were unsuccessful, the Gunn oscillator was packaged in a high-Q waveguide cavity to improve its spectral purity to a point

where it does not degrade the paramp noise performance. The Gunn oscillator is followed by a ferrite isolator and a waveguide-to-miniature coaxial adapter. The oscillator operates on primary power consisting of 8 Vdc at 0.5 A, making it readily adaptable to spacecraft power sources.

Table 1 presents some of the test results measured on the paramp. The 1-dB bandwidth of 90 MHz at 15-dB gain falls just 10 MHz short of the desired coverage of 2.2 to 2.3 GHz. The noise temperature measured 139° K, better than the design specification of 150° K. After a brief warm-up, the 24-hour gain stability was within ±0.2 dB under laboratory conditions, which indicates excellent inherent long-term gain stability. The one area that is deficient is the operating input VSWR. This quantity, although higher than operational requirements allow, is a fairly easily correctable fault. Since it is due to the isolation of the input circulation being too low, it can be improved in future efforts by better matching and optimization of the circulator-paramp physical layout and by the inclusion of an output isolator in the integrated package.

Table 1

Integrated Paramp Test Data

Gain	Bandwidth (1 dB)	Noise Temperature	Gain Stability (24-hour)	Input VSWR	Weight	Volume
15 dB	90 MHz	139° K (1.7 dB)	±0.2 dB	4:1	9 oz	6 $in.^3$

The overall weight of the unit is 9 oz, and the overall volume is 6 in^3. The weight and volume figures can be reduced significantly by using lighter weight materials in the pump circuitry, and designing special waveguide components for the pump circuit instead of using standard flanged components. The dc input power requirements of 4 W can also be lowered to less than 1 W by the use of more efficient Gunn devices.

This integrated paramp is the only broadband, low-noise, stable unit successfully developed to date. It is felt that, with some further improvement in several deficient areas, a truly operational integrated paramp will be available for use in various spacecraft systems as well as in ground applications when extremely small weight and volume are critical requirements. Also, a program to extend these accomplishments to the 15-GHz communication band is now in progress.

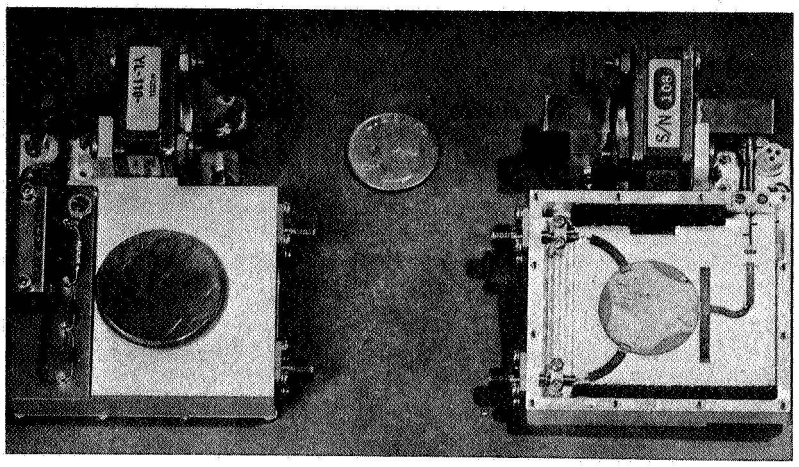


Figure 1

PRELIMINARY ANTENNA RADIATION PATTERN CALCULATIONS FOR THE ATS F FLEX-RIB REFLECTOR

Richard F. Schmidt

CHAIRMAN:

Our next speaker is Mr. Richard Schmidt, who will discuss preliminary calculations of the antenna radiation patterns for the ATS F reflector.

MR. SCHMIDT:

For the past two or three years, we have been developing a general-purpose computer program to provide radiation patterns for a variety of reflector-antenna configurations. The program, based on the Kirchhoff approach, uses vector analysis and differential geometry to determine complex-vector sheet currents on a reflector surface. Adequate sampling of the sheet currents permits numerical integration and evaluation of the back-scattered radiation fields in space. High-speed digital computers, such as the IBM 360/91, are used to evaluate the otherwise intractable integrals arising from the Kirchhoff method.

Since the full-scale (30-ft) ATS F reflector will consist of 48 supporting ribs and stretched, metallized Dacron mesh, it will be a very close approximation to an ideal paraboloid. An attempt to confirm computed results by means of a full-scale measurement using a 30-ft reflector would be undesirable since fabrication costs would be high and the resulting gain degradation would be very small and difficult to measure. Program verification was obtained via preliminary calculations and measurements based on a 4-ft, 12-panel reflector antenna (Figure 1).

Figure 2 shows a composite of the computed data for the 4-ft, 12-panel ATS F test model antenna and the ideal reference paraboloid with F/D = 0.375. These are not predictions for the ATS F full-scale antenna. Results for the -11.5-dB edge-taper illumination of the reflector, as well as the -16.0-dB edge-taper results, are shown. Edge-taper is controlled by the shape of the prime-feed pattern and choice of feed point. Since the illumination distribution is constant, side-lobe levels are constant over the frequency band for the ideal paraboloid. An explanation of the reduced first side-lobe levels of the test antenna is not straightforward since both amplitude and phase distributions are different from those of the ideal paraboloid.

Beamwidth information, obtained from the same computations, is shown directly below the side-lobe level data. As expected, the panel effect is very slight at the lower frequencies. Beamwidth-broadening gradually increases with increasing frequency until, at 16 GHz, it is approximately 12% of that for an ideal paraboloid.

Figure 3 shows the computed gain loss of the ATS F 4-ft, 12-panel test model antenna when compared to an ideal 4-ft paraboloid. Again, there are curves for both the -11.5-dB and the -16.0-dB edge-tapers. A rapid loss of performance, about 4 dB at 16 GHz, is observed with increasing frequency. Again, this is for the test model antenna used to verify the computer program; no inference is made concerning the full-scale ATS F antenna performance.

Measured data for gain degradation are superimposed on the preceding graph. Experimental edge-taper values ranged between -10.4 dB and -10.7 dB, leading to gain degradation values of -0.35 dB at 4 GHz, -1.2 dB at 8 GHz, and -1.8 dB at 12 GHz. The tolerance on this measured data is at least ±0.1 dB. In view of the smaller edge-taper, it can be seen that the agreement is very good at 4 and 8 GHz. The measured data at 12 GHz should lie below the -11.5-dB edge-taper plot obtained by computer, but it does not.

In conclusion, the agreement between the Kirchhoff approach and actual measurement is good for the flex-rib design. Such data comparison is performed on a continuing basis. It is noted that, due to gravitational effects, radiation-pattern measurement of space-qualified antennas is difficult in the Earth environment; a proven mathematical simulation is, therefore, an especially valuable tool. The effects of thermal distortion of reflector surfaces on parameters such as gain and beamwidth have also been calculated by this method. Preliminary programs have been written, and tested, to predict monopulse null-shift and other antenna characteristics over the entire satellite orbit.

CHAIRMAN:

Are there questions for Mr. Schmidt?

MEMBER OF THE AUDIENCE:

Is the pilot model generally going to be the type that is on ATS G also?

CHAIRMAN:

Yes, there will be a 30-ft dish antenna on the ATS G.

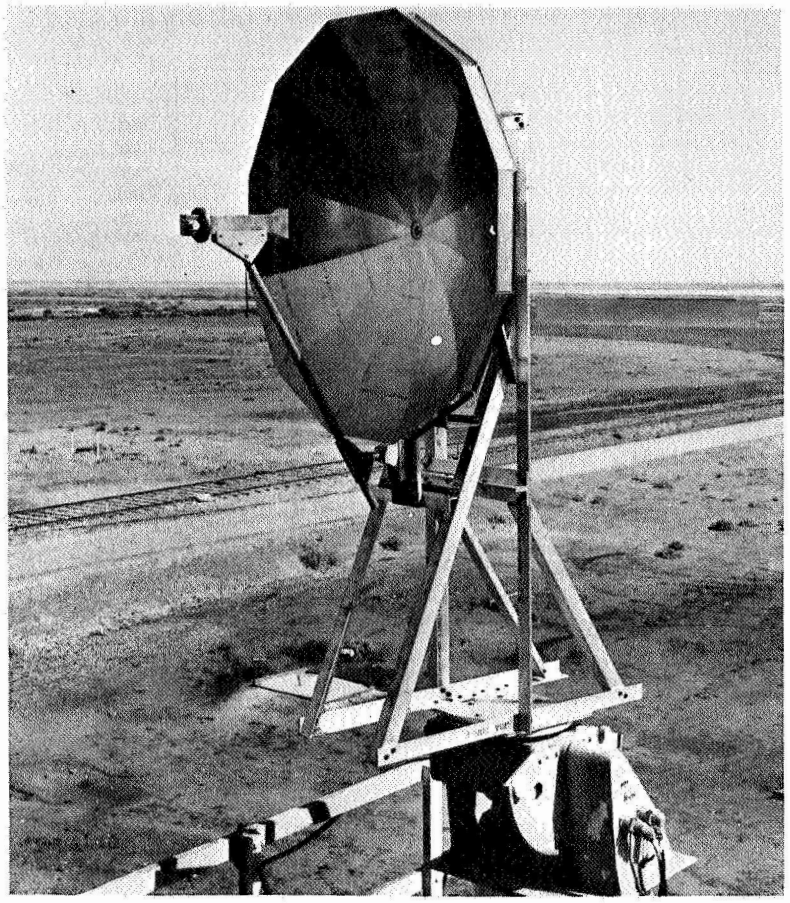


Figure 1–ATS F test model (4 ft, 1 2 rib).

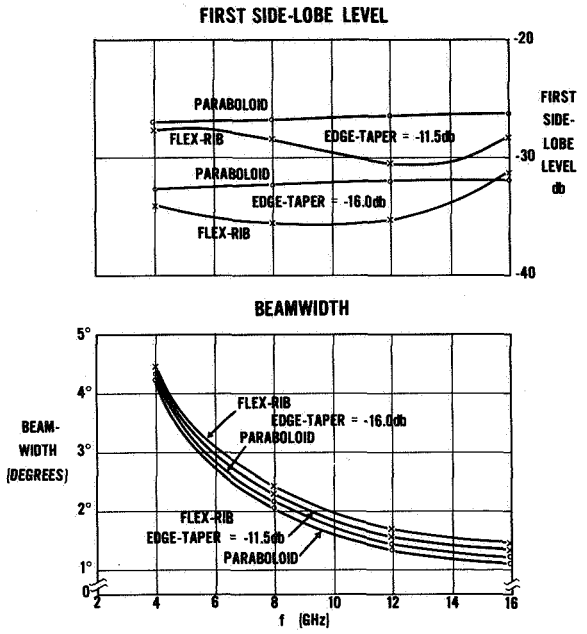


Figure 2–ATS F test model beamwidth and side-lobe level versus frequency.

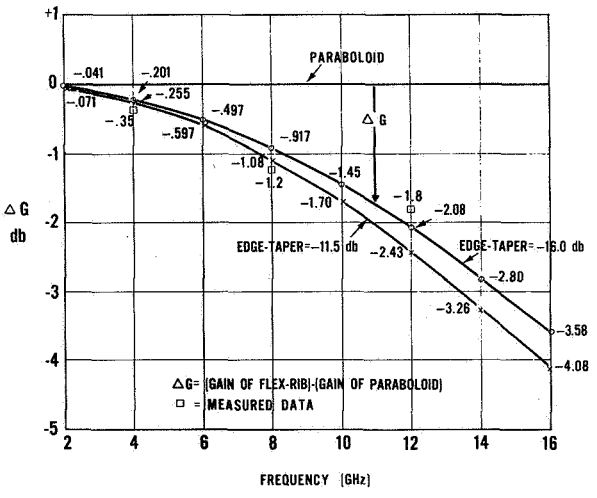


Figure 3–ATS F test model gain loss due to panels versus frequency.

LASER PROPAGATION THROUGH ATMOSPHERIC TURBULENCE

Michael W. Fitzmaurice

CHAIRMAN:

Mr. Michael Fitzmaurice will discuss laser propagation through atmospheric turbulence.

MR. FITZMAURICE:

The efficient use of lasers for communication through the Earth's atmosphere requires a detailed understanding of the effects of turbulence upon the beam. Analyses have gradually evolved to the extent that today we have available a nearly complete theoretical description of the amplitude and phase fluctuations throughout the beam. The present situation, then, is that we must test the theory and determine its validity, or its range of validity. If indeed the theoretical predictions are accurate, then atmospheric laser communication systems can be designed with confidence of optimum performance.

To conduct propagation experiments, we have instrumented a horizontal range at the GSFC optical research facility. The range is 0.6 km long, and the optical path is folded and extended to 1.2 km by means of a 51-cm flat mirror. A horizontal test range has two primary advantages. First, one can expect the statistical properties of the turbulence to be constant over the path; this minimizes the number of meteorological stations required and permits experimental results to be compared to theory in a straightforward manner. Second, one can instrument a near-ground horizontal range much more easily than any other type of range.

One of the more important relationships predicted by theory is the dependence of laser-beam amplitude fluctuations on laser wavelengths. Considering the specific case of an He-Ne laser ($.633\mu$) and a CO_2 laser (10.6μ), wave theory predicts that the CO_2 fluctuations will be reduced by a factor of 26.7 compared to the He-Ne fluctuations. We have conducted an experiment to test the accuracy of this prediction. The experimental schematic is shown in Figure 1. An He-Ne and a CO_2 laser were aligned parallel and transmitted over the 1.2-km folded path. The return beams were detected by a photomultiplier and a gold-doped Ge detector; a portion of the CO_2 output was picked off by a beamsplitter and recorded so that internal laser power fluctuations could be monitored. The recorded data were later played back through an A/D converter into a digital computer. The computer analyzed the statistics of the amplitude fluctuations; the results are shown in Figure 2. Along the y-axis, the quantity ψ is the ratio of the variance of the He-Ne

fluctuations to the variance of the CO_2 fluctuations. The strength of turbulence, as characterized by the refractive index structure constant C_N^2, is plotted along the x-axis. The value 26.7 is the theoretical prediction, and the average of the 39 independent data points measured is 26.8. Clearly, the agreement with theory is extremely good. The computer also analyzed the probability density function of the data and found it to be log-normal for both wavelengths. In terms of applications, this result says that in a spacecraft-to-ground laser system, such as the ATS F system, one can predict the signal-fading characteristics of the infrared laser signal based on visible stellar scintillation measurements.

One of the methods proposed for avoiding the problem of atmospheric turbulence in laser communication systems is the idea of spectral diversity. This technique requires the transmitter station to operate at more than one wavelength in the hope that, at the receiver, the outputs of the tuned detectors can be combined so as to reduce the atmospheric noise. The actual utility of this technique depends heavily upon the degree of correlation of the atmospheric noise imposed on the various transmitted beams. Figure 3 shows the experimental setup that was designed to measure the correlation between different wavelength laser beams which have traversed the same path. Two continuous-wave gas lasers (He-Ne and Ar) were positioned parallel, and the output beams were superimposed by a mirror-beamsplitter arrangement. Each beam had a divergence of approximately 1 mrad and a diameter of 3 mm. The combined beams were then transmitted over the 1.2-km range. The return energy was detected through the same 3-mm entrance aperture by two photomultipliers which were equipped with optical filters to reject unwanted wavelengths. The He-Ne laser was operated at 6328 Å, and the Ar unit radiated at any one of eight wavelengths from 5145 Å to 4579 Å. The detector outputs were processed through analog circuitry to obtain the spectral correlation coefficient for the particular wavelength separations available.

The data shown in Figure 4 were obtained during both strong (mid-afternoon) and weak (night) turbulence conditions and are shown with a least-squares linear fit. Clearly the spectral correlation function depends on the strength of turbulence, but the variation with wavelength separation is not detectable over the 600 Å range measured. On the basis of the data, it appears that spectral diversity offers little help unless sources with wavelength separations of 1μ or more are used.

In addition to the horizontal range experiments, we are now preparing an experiment that, using balloons, will obtain data over a vertical propagation path. We expect to do this experiment during the summer of 1970.

MEMBER OF THE AUDIENCE:

I noticed in the schematic for your first experiment that you processed and chopped the 10.6μ radiation but not the other. Is that the correct portrayal, and if it is, what is the reason for that?

MR. FITZMAURICE:

Yes, it is correct. There are a couple of reasons, really. One basic reason is that the 10μ detectors do not operate well at dc at baseband. The one we used operated with a very heavy bias, and in addition, the noise characteristics are better if you can get away from dc. So, you simply move the spectral density of this phenomena up to a higher frequency where you can look at it in a less noisy environment.

MEMBER OF THE AUDIENCE:

Were you able to take out the effect of modulation losses, chopping, and demodulating in determining the scatter?

MR. FITZMAURICE:

Well, you see, we were not interested in average transmission. What we were looking at were the statistics of very fast but large fluctuations. So we were not concerned with the fact that you lose perhaps 50% of your power by chopping the beam on and off. That was not the purpose at all. The purpose was to look at the statistics of the atmospheric-induced fluctuations, not average transmission.

MEMBER OF THE AUDIENCE:

The variation of statistical properties with time and environmental conditions, as well as the dynamic range of atmospheric disturbances, is extreme, particularly as you go to longer range. Have you done any tests to determine the statistical accuracy of your results?

MR. FITZMAURICE:

At the time we did this work, we did not have any way of directly, you might say unambiguously, looking at the atmosphere, such as with thermal sensors. We do now.

As far as dynamic range goes, of course the CO_2 fluctuations are characteristically weak. There is no dynamic range problem. Our dynamic range is about 46 dB, which, based on analyses which are just coming out now, seems to indicate that we could have handled log-amplitude variances

of about 0.30 to 0.35. We never got that high. We do not think that was a significant problem in this case. It should be kept in mind that these were nighttime data.

MEMBER OF THE AUDIENCE:

I would like to add some comments to that. We have had different experience in a different program. MSFC is running a Doppler velocimeter to measure winds.

MR. FITZMAURICE:

To measure winds?

MEMBER OF THE AUDIENCE:

Yes. They have done it and we have just finished conducting a joint program with ESSA. The results are being analyzed by very sophisticated correlation techniques.

MR. FITZMAURICE:

Yes, I am familiar with the correlation technique they use at ESSA; the output of their analysis depends upon the fact that you had Gaussian random variables to start with. I back off immediately from that. I do not like that assumption. It may be true, but I do not like the additional assumption.

MEMBER OF THE AUDIENCE:

They analyze in a different way now. But what I am saying is that these experimental results are available, and I am not so sure that the preliminary groups are quite correct.

MR. FITZMAURICE:

Well, we are doing more work. We are taking additional data here too, and we do not see anything very different from what we had before. There may be some minor modifications. This will be published.

If you are speaking of the spatial spectrum of the turbulence itself, all I can say is that when we did take that data, at that time we did not have a handle on it. We had no way of even looking at it. It was simply a very optical-type technique. There were no thermal sensors involved. So whether the spatial spectrum was Kolmogoroff or something else, I really do not know.

I would have to say that we now look at it closely. We do correlations on the time spectrum, and you can infer from that what the spatial spectrum was, if you believe Taylor's hypothesis.

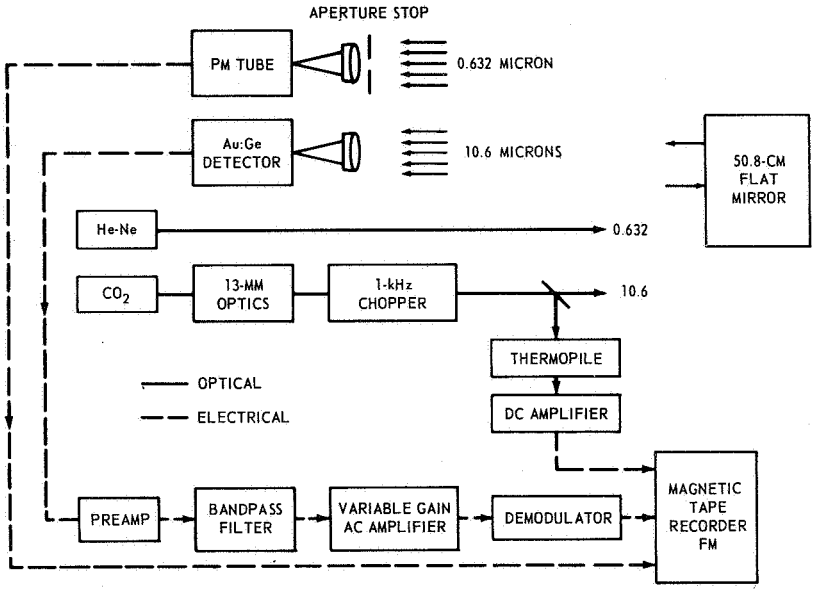


Figure 1

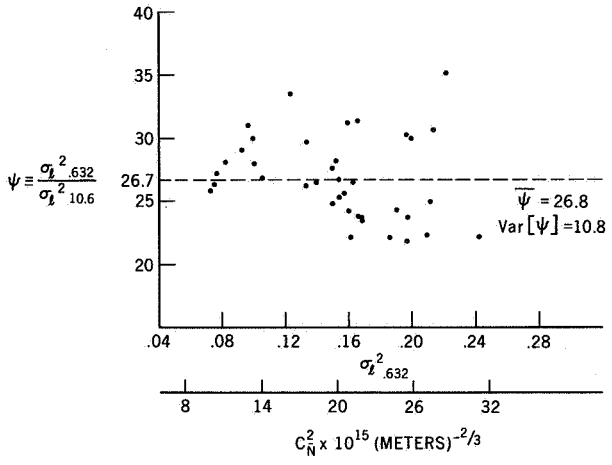


Figure 2

EXPERIMENTAL SCHEMATIC

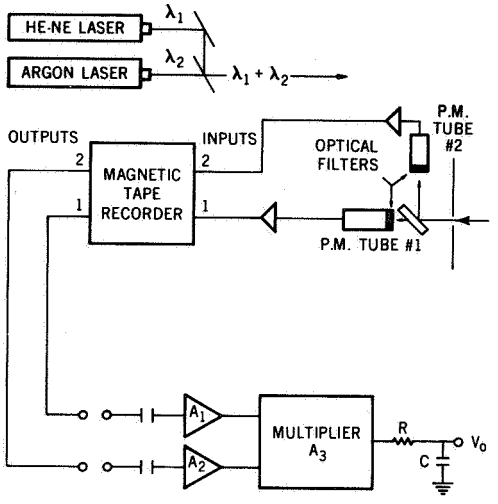


Figure 3

SPECTRAL CORRELATION DATA

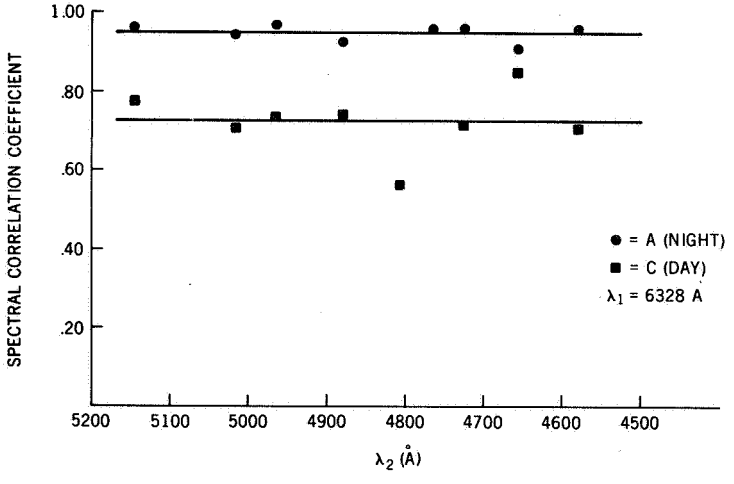


Figure 4

10.6-MICRON LASER STABILIZATION

William E. Richards

CHAIRMAN:

Our next speaker, Mr. William Richards, will discuss 10.6μ laser stabilization.

MR. RICHARDS:

If you were told that we had achieved absolute CO_2 frequency stabilities of parts in 10^9, it probably would not sound as impressive as being told that we had learned to control the optical or electrical path length to better than 12 Å, or 10^{-9} meters. In fact, both statements are identical and represent one of our accomplishments for the past year. Furthermore, we have performed heterodyne experiments and achieved heterodyne beat-frequency stabilities of parts in 10^{11}. We have achieved these stabilities by applying RF stabilization techniques to laser technology and by using the CO_2 molecular emission profile as a frequency reference.

Figure 1 illustrates the power-frequency profile for three of the stabilization techniques that are used. They are (1) dithering to line center (LCD), (2) locking to the laser power profile (AGC), and (3) automatic frequency control (AFC).

The LCD technique applies a signal from a reference oscillator to the laser and dithers its output over the laser profile peak. The laser output is detected and compared with the reference oscillator signal in a phase-lock amplifier. Any error signal is then fed back to the laser. The AGC technique monitors the output power of the laser and compares the laser detector voltage variations with a reference dc source in a differential amplifier. Error signals from the differential amplifier are fed back to the laser. These two techniques use the laser profile as a reference.

The AFC technique monitors the laser heterodyne beat and uses a discriminator for a reference. The physical length of the laser varies when the error signals are applied to a piezoelectric crystal upon which one of the laser mirrors is mounted.

In order to analyze the laser noise spectrum and relative effectiveness of each stabilization loop, we monitored the heterodyne beat output with a second-order phase-lock loop tracking filter. Although this is a standard RF tool, it is a welcome addition to optical analysis since it not only

gives a measure of the intensity of any frequency noise but also gives the frequency of the noise, thus providing identification of the noise source.

When analyzing the results of the above experiments, we discovered two other sources of frequency instability that had previously been ignored, namely, atmospheric- and gravitational-induced cavity length variation. For example, in the transition from an Earth to a space environment, a frequency offset of about 60 MHz would result for a CO_2 laser; although this is not catastrophic, it must be compensated for in initial laboratory calibrations.

One of the difficulties with an internal laser modulator is that its index of refraction changes greatly with temperature. This, in turn, causes large laser frequency excursions. To date, the only method used to circumvent this instability is modulator temperature control to 0.1° or 0.01°C. Figure 2 illustrates a new technique for passive modulator stabilization which we have proposed and experimentally verified. This technique is based upon recognizing the existence of a low-loss, highly transmissive, readily available material with a large negative change in index of refraction with temperature. This material is KRS-5, or thallium bromide iodide. The properties of the material are such that proper selection of its relative length with respect to the length of the modulator crystal will theoretically eliminate frequency deviations resulting from modulator temperature changes. Practically, minor frequency variations over ±10°C seem easily achievable. Using KRS-5, this technique is applicable not only at 10.6μ but also from the visible to far infrared region.

In conclusion, a summary of our achievements for the past year includes–

1. Absolute frequency stability of 3 parts in 10^9.

2. Heterodyne frequency beat stability of 3 parts in 10^{10} (3 parts in 10^{11} with new supply).

3. Use of ϕ-lock loop tracking filter as optical analytical tool.

4. Demonstrated importance of the effects of pressure and gravity on laser stability.

5. Proposed and experimentally verified dn/dT frequency control (CO_2, solid state frequency references).

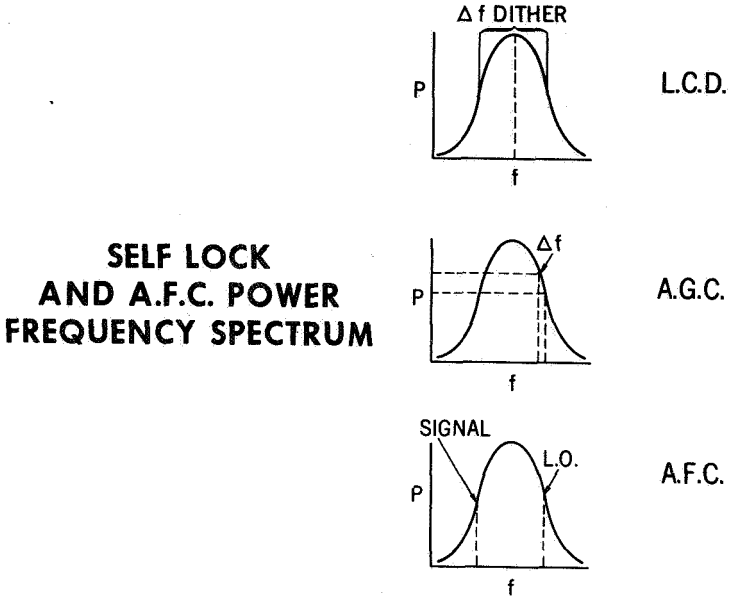


Figure 1

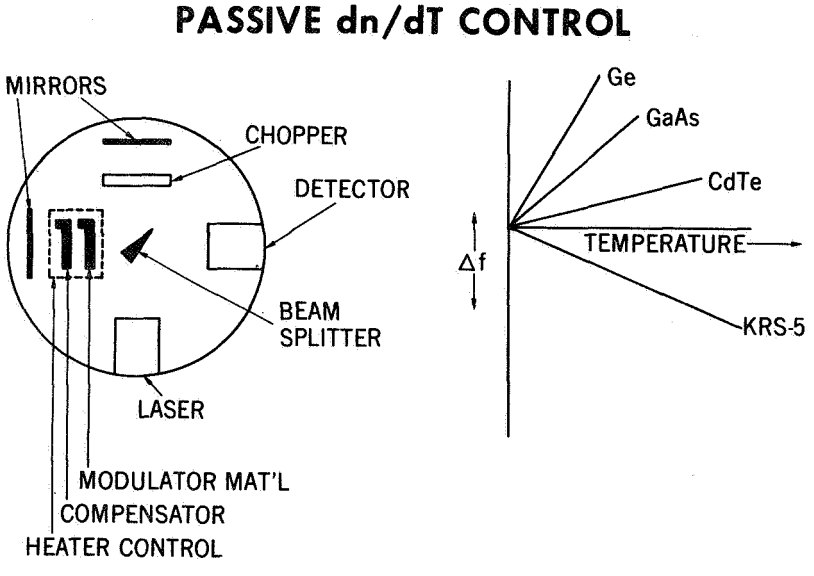


Figure 2

EXPERIMENTAL ANALYSIS OF MERCURY CADMIUM TELLURIDE PHOTOVOLTAIC IR MIXERS FOR SPACECRAFT RECEIVERS

John H. McElroy

CHAIRMAN:

Our next speaker is Mr. John McElroy, who will discuss the experimental analysis of mercury cadmium telluride photovoltaic IR mixers for spacecraft receivers.

MR. McELROY:

In 1966 work was begun on a program to develop wideband receivers for the CO_2 laser. The development program has advanced the frequency response of 10.6μ IR mixers from 1 MHz at an operating temperature of 4.2° K in 1966 to 1.2 GHz at the same operating temperature in 1968. Although a 4.2° K operating temperature is not inappropriate for ground stations, it would rule out use in a spacecraft system. Therefore, during the past year, efforts were begun to develop mixers that could operate at temperatures of 100° K or higher with the same wide frequency response. An operating temperature in excess of 100° K permits the use of lightweight, completely passive, radiation coolers with indefinitely long operating lifetimes. The purpose of this work is to develop CO_2 laser communication systems that would be applicable in spaceborne systems such as the Tracking and Data Relay Satellites and Earth Resources Satellites.

The most appropriate mixer material for 100° K operation is the semiconductor, mercury cadmium telluride. Three forms of this semiconductor have been examined: n-type, p-type, and p-n junctions. Of these, only the p-n junction, used as a reverse-biased photodiode, has shown the positive potential for frequency response beyond 1 GHz.

As an intermediate goal, a mixer was developed for the ATS F laser communication experiment. This experiment requires a mixer having flat response to 35 MHz. It also requires that the total thermal dissipation in the mixer, due to both local oscillator and bias powers, be sufficiently small to permit the use of a radiation cooler. This intermediate goal has been achieved and surpassed.

The experimental results which will now be presented were obtained from a specially constructed photodiode in which care was taken to minimize parasitic inductance and capacitance, to minimize junction capacitance, and to maximize reverse resistance.

Figure 1 shows the minimum detectable signal power as a function of local oscillator power. It can be seen that increasing the local oscillator power from 200 μW to 2 mW causes the minimum detectable signal to decrease asymptotically to 2.08 x 10^{-19} W/Hz. Since these measurements were made at an operating temperature of 77° K on a mixer element designed for peak performance at 100° K, somewhat better results should be obtained at the designed operating temperature. It now appears that a minimum detectable signal of less than 10^{-19} W/Hz at 100° K is attainable.

Figure 2 shows the minimum detectable signal power as a function of intermediate frequency. It can be seen that the mixer response is flat to approximately 40 MHz. The total power dissipated in the mixer is approximately 1 mW.

This is the first report of a near-quantum-limit performance by a mercury cadmium telluride photodiode and the first report of a 100° K IR mixer having sufficient frequency response to meet the ATS F laser communication experiment requirements. In our latest experiments, we have obtained samples with flat responses beyond 100 MHz at 100° K.

MEASURED P_{MDS} FOR PHOTOVOLTAIC $Hg_{1-X}Cd_XTe$ AT 10.6 MICROMETERS

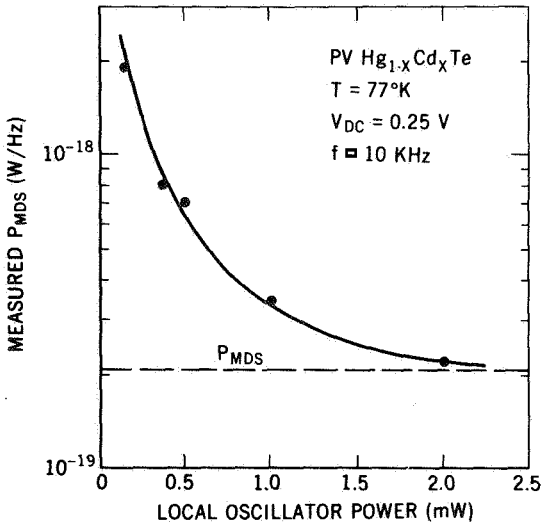


Figure 1

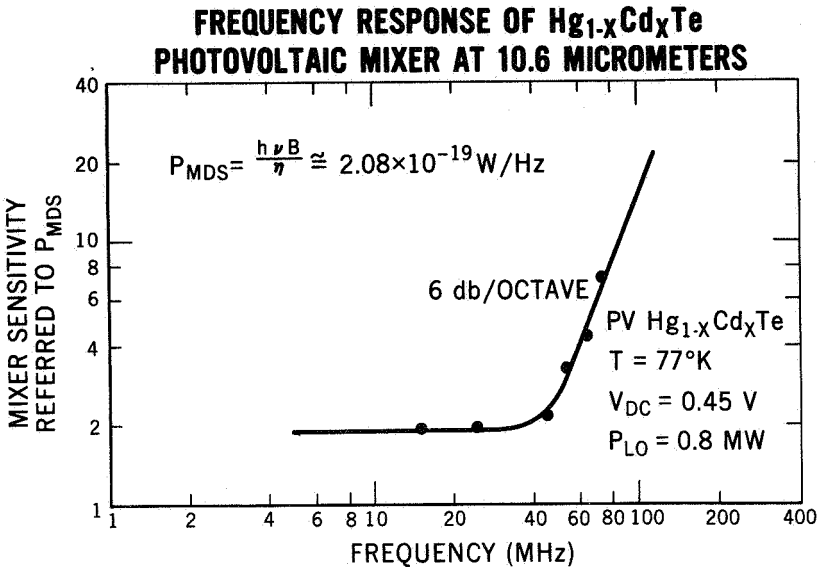

FREQUENCY RESPONSE OF $Hg_{1-X}Cd_XTe$
PHOTOVOLTAIC MIXER AT 10.6 MICROMETERS
$P_{MDS}= \frac{h\nu B}{\eta} \cong 2.08\times10^{-19}$ W/Hz
6 db/OCTAVE
PV $Hg_{1-X}Cd_XTe$
T = 77°K
V_{DC} = 0.45 V
P_{LO} = 0.8 MW
MIXER SENSITIVITY REFERRED TO P_{MDS}
40
20
10
8
6
4
2
1
1
2
4
6
8
10
20
40
60
80
100
200
400
FREQUENCY (MHz)

Figure 2

USE OF COHERENT OPTICS FOR ON-BOARD DATA PROCESSING

James P. Strong III

CHAIRMAN:

The next series of papers will be on data handling. Mr. James Strong will discuss the use of coherent optics for on-board data processing.

MR. STRONG:

Coherent optics may well provide an easy method for doing involved computations on board a spacecraft. An immediate result of such computations would be a substantial reduction in the amount of information that has to be sent back from a spacecraft.

As an example, consider a mission to Jupiter where we are interested in studying cloud formations and motions. We would probably have only enough transmitter power to send back a few Tiros or Nimbus type pictures. However, using coherent optics, we can also process a large number of cloud cover pictures on board and send back such words as "structured," "nonstructured," "small cloud formations," or "large cloud formations."

To see how this can be done, consider Figure 1, which shows two cloud cover pictures next to their Fourier transforms. By measuring the distance information lies from the center of the Fourier transform, one can tell whether a picture is made up mostly of small scattered clouds or mostly of large clouds. Large cloud masses (Figure 1a) have Fourier transforms that are more concentrated toward the center (Figure 1b) than the Fourier transforms of small, scattered clouds. Clouds that have frontal patterns such as in Figure 2a produce information concentrated along a line in the Fourier transform plane, as seen in Figure 2b.

In order to detect this type of information we have developed two detectors diagramed in Figure 3. The patterns are etched on a silicon solar cell. The rings in the radius detector measure the intensity of light at various distances from the center of the Fourier transform; the sectors in the angle detector detect concentrations of light along a line at various angles. For example, Figure 4 shows two cloud cover pictures: one with scattered clouds, the other an overcast system with parallel line structure. The graphs below each picture show the outputs of the detectors when placed in the Fourier transform plane.

One can easily see the difference in these outputs for the two pictures. To send back these graphs would require about 150 bits of information. To send back a 128- by 128-element binary representation of the original picture for classification on the ground would require about 16,500 bits. Thus, the quantity of information sent back has been reduced by a factor of more than 100. With this detector then, we have a sensor that will produce electrical signals representing information about the Fourier transform of a picture, which can provide data from an optical system to the input of an on-board digital system.

CHAIRMAN:

Are there any questions of Mr. Strong? I see quite a few hands here.

MEMBER OF THE AUDIENCE:

I am wondering how you make these characterizations of the transforms of these pictures?

MR. STRONG:

The characterizations are not unique. All we are trying to do is determine whether any one picture is made up of small clouds or of large clouds, or whether there is any kind of structure at all. Any picture with that particular type of information in it will give us basically the same type of Fourier transform.

MEMBER OF THE AUDIENCE:

Can meteorologists use this kind of information?

MR. STRONG:

We have often suggested this, but they seem much more inclined to get the pictures themselves.

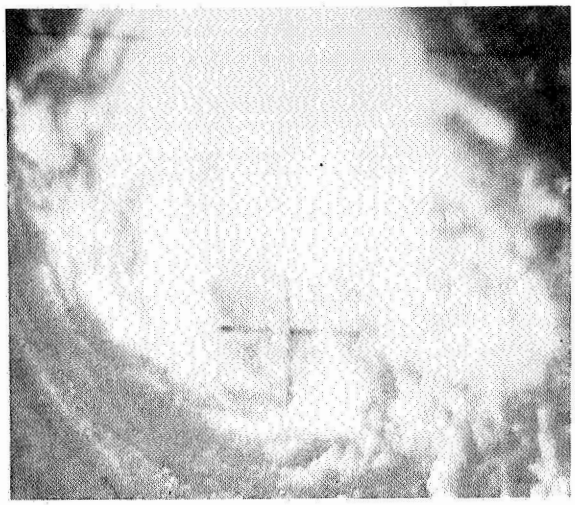


Figure 1a

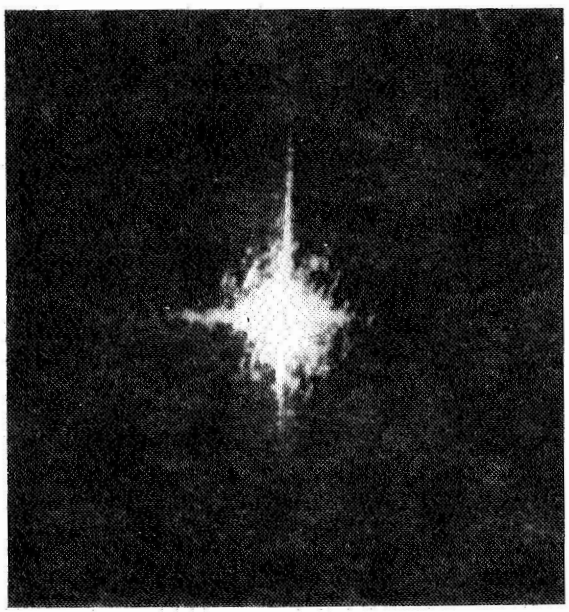


Figure 1b

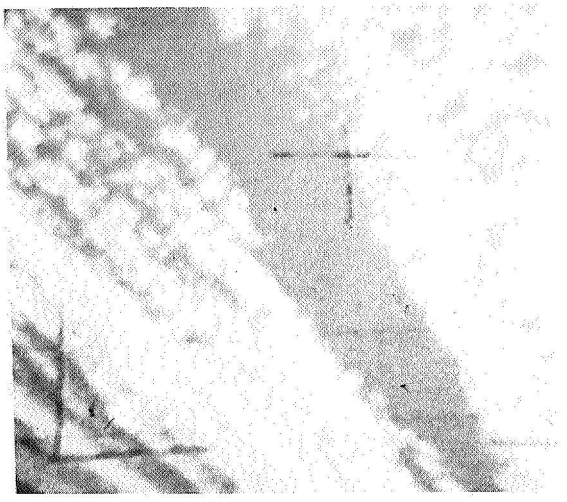


Figure 2a

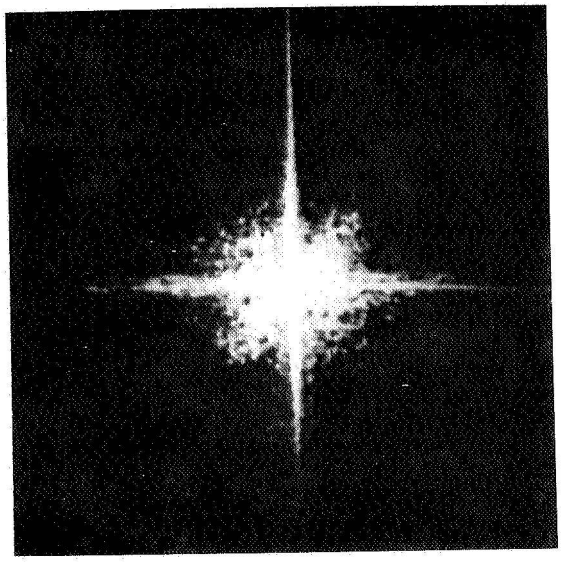
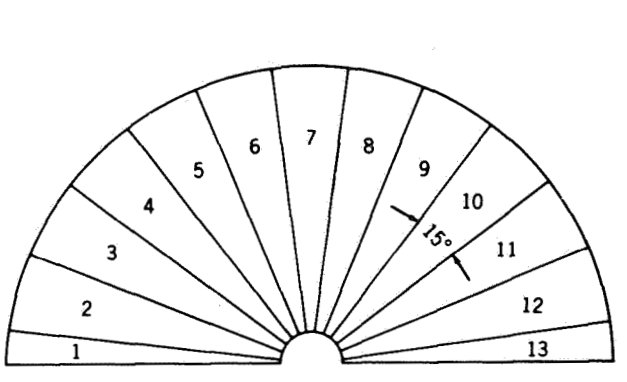
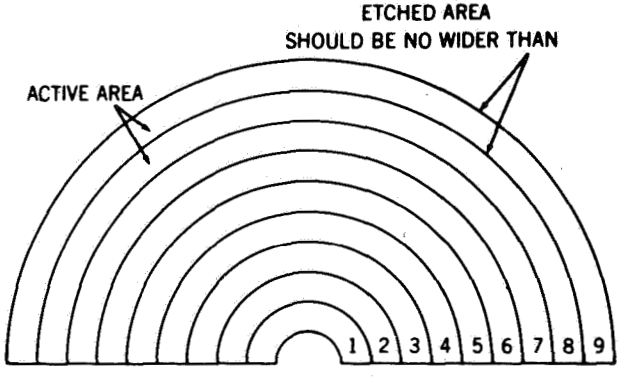


Figure 2b



ANGLE DETECTOR



RADIUS DETECTOR

Figure 3

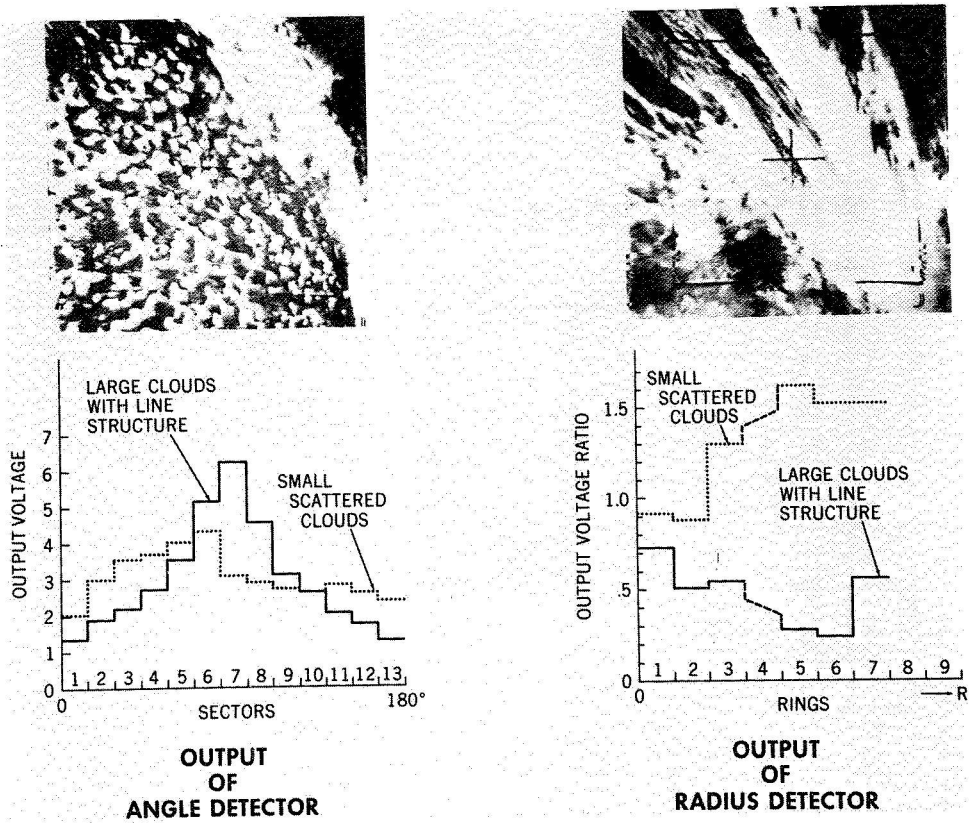

LARGE CLOUDS
WITH LINE
STRUCTURE
SMALL
SCATTERED
CLOUDS
OUTPUT VOLTAGE
0
1
2
3
4
5
6
7
1 2 3 4 5 6 7 8 9 10 11 12 13
0
SECTORS
180°
OUTPUT
OF
ANGLE DETECTOR
SMALL
SCATTERED
CLOUDS
LARGE CLOUDS
WITH LINE
STRUCTURE
OUTPUT VOLTAGE RATIO
0
.5
1.0
1.5
1 2 3 4 5 6 7 8 9
0
RINGS
R
OUTPUT
OF
RADIUS DETECTOR

Figure 4

ADVANCES IN COMPLEMENTARY MOSFET TECHNOLOGY

David H. Schaefer

CHAIRMAN:

We go now to Mr. Dave Schaefer, who will describe advances in complementary MOSFET technology.

MR. SCHAEFER:

The development of complementary metal-oxide-silicon-field-effect-transistor (MOSFET) technology will lead directly to the realization of low power, light weight, sophisticated data systems--systems that are needed for everything from space shuttles to Mercury probes.

MOSFET's are like vacuum tubes in that they have no grid or gate current. They are also easy to make. Large scale integrated (LSI) units have been fabricated with up to 3000 devices on a chip. Complementary MOSFET technology employs both p- and n-type MOSFET's connected such that there is never dc current flow from the power supply to the ground.

Figure 1 is a picture of a silicon chip that contains both n- and p-type MOSFET's and npn and pnp bipolar transistors. The reason for attempting to make MOSFET's and bipolar transistors on the same piece of silicon is that MOSFET's are better suited for the logic needed for LSI circuits. To quickly charge large interpackage capacities, however, requires the driving of large currents. If only MOSFET technology is used, the current drivers require units that take up a great deal of area. On the other hand, bipolar transistors, for the currents we have under consideration, need only one-tenth of the area of their MOSFET counterparts. Silicon real estate is very expensive from a yield point of view. Therefore, we envision large LSI circuits that have MOSFET interiors with a periphery of bipolars that communicate with the outside world.

There is a different method of breaking out of our real estate problem. Figure 2 shows a circuit board upon which is a piece of aluminum foil. On the aluminum foil are three deposited complementary MOSFET inverters. This technology utilizes tellurium for the p-type devices and cadmium selenide for the n-type devices. It is our hope that this technology can be developed so that very long shift registers can be put on rolls of substrates such as aluminum foil. This would lead to a new type of mass storage medium.

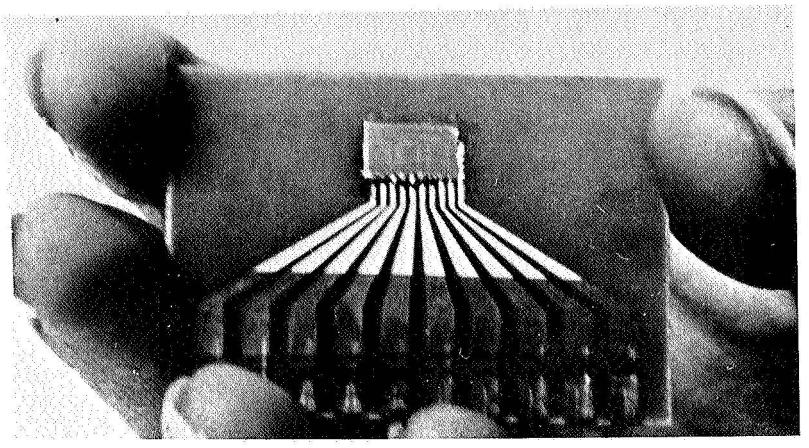


Figure 2

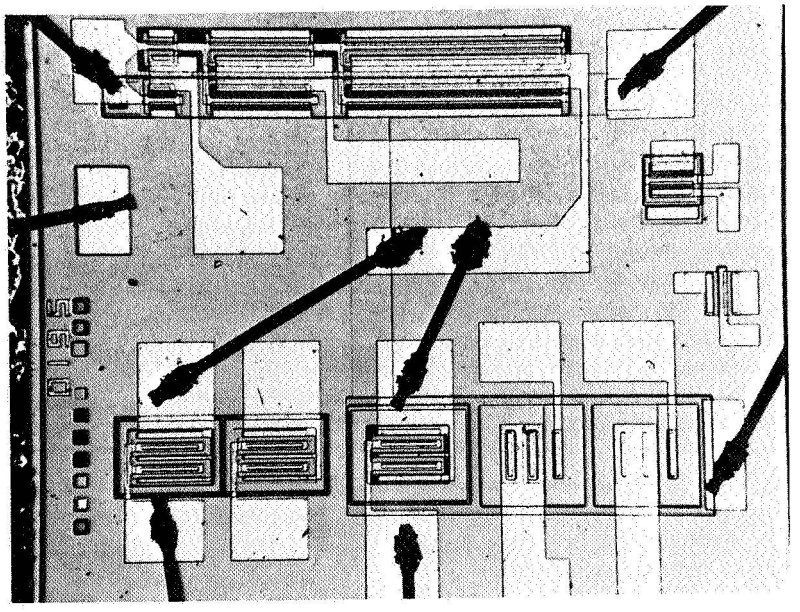


Figure 1

LOW FLUX LEVEL NEUTRON EFFECTS ON MOSFETS

Phillip A. Newman

CHAIRMAN:

The next paper will be presented by Mr. Phillip Newman. Its title is, "Low Flux Level Neutron Effects on MOSFET's."

MR. NEWMAN:

Until recently, NASA has not been concerned with the effects of neutrons on spacecraft electronics since neutrons are virtually nonexistent in the space environment. In the past few years, however, two developments have made a closer look at this necessary.

The first is the proposed use of radioisotope thermoelectric generator (RTG) power supplies on deep space missions which, with their neutron and gamma fluxes, provide a long term hostile environment for electronic systems. Second is the increasing use of the somewhat radiation-sensitive MOSFET in spacecraft electronics. The gate drain characteristic of the MOSFET shifts under ionizing radiation, tending to turn p-channel devices off and n-channel devices on. The radiation induced shift is also strongly dependent on the gate bias during irradiation.

Since the effects of gamma radiation on MOSFET's have been studied and are reasonably well understood, it was decided to attempt to isolate the effects of neutron radiation. Prior work used reactor neutrons as a radiation source. The results observed can be mostly attributed to the gamma radiation background present in all reactor spectra. For this experiment (shown in Figure 1), an available low yield Pu^{239}-Be source was used. While the neutron spectrum of this source is very close to that of the RTG, it has only one gamma photon for each 1000 neutrons, compared to 35 gamma photons for each neutron in the RTG.

Since only eight devices could be accommodated conveniently in the experiment, four off-the-shelf commercial MOSFET's from three manufacturers were chosen. The devices were biased at -10 V, and the gate threshold voltage was measured at 12-hour intervals. The experiment ran for 230 days.

A typical plot of the experimental results (Figure 2) shows that the neutron-induced shift in the gate threshold voltage fits a power law of the form

$$dV = K(\text{dose})^{0.76}.$$

All of the devices used in this experiment fit this type of law quite closely. The total shift in gate threshold voltage to date is shown to be close to 1 V. The total accumulated dose of neutrons for a typical 5-year deep space mission that used an RTG power supply has been estimated to be 4×10^{10} neutrons/cm^2.

On the basis of this data, a shift of about 50 mV due to neutrons can be expected in the gate threshold voltage of a MOSFET during the 5-year mission. It is apparent that a shift of this magnitude would not seriously affect most electronic systems that used MOSFET's. In addition, the power law dependence observed will make it possible to forecast with some degree of confidence the neutron degradation for other configurations of the RTG power supply.

It should be emphasized that this work has covered the effects of neutrons only. For proposed RTG configurations there will be a sufficient flux of gamma rays to cause equal or greater damage to the MOSFET.

CHAIRMAN:

Any questions for Mr. Newman?

DR. PIEPER:

The curve in Figure 2 is based on results from a dozen devices; what is the spread in the individual devices relative to the size of the values plotted?

MR. NEWMAN:

At the extreme right, the curve shows approximately a 1-V shift. The spread might be 0.5 V altogether from largest ΔV to smallest ΔV at this fluence.

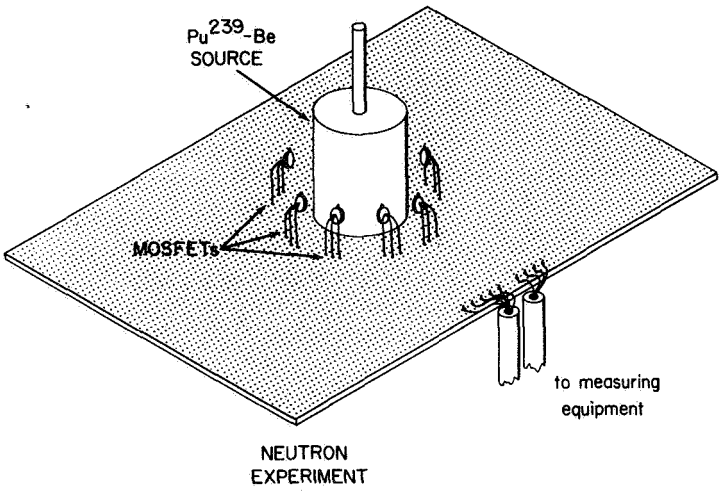


Figure 1

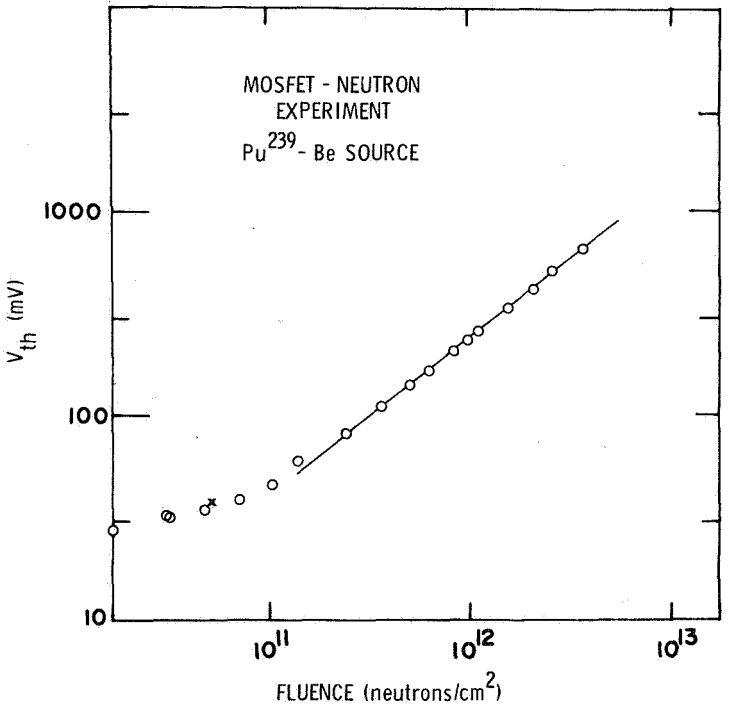


Figure 2

SOLID STATE SHIFT REGISTER AS A DATA STORAGE DEVICE

John C. Lyons

CHAIRMAN:

The next speaker is Mr. John Lyons, and the topic of his paper is the solid state shift register as a data storage device.

MR. LYONS:

There is a need for small data storage capabilities in a great many experiments on scientific spacecraft. Often this requirement does not exceed 1,000 bits of storage. For such a small data storage need, it is not efficient to attempt to utilize a tape recorder or a magnetic core memory and the circuitry required to drive it. The weight and power required for a tape recorder make it impractical to use such a device for less than 50,000 bits of memory. The large amount of control circuitry required to drive a magnetic core array also makes it impractical to utilize this approach for less than 4,000 bits of memory. For this reason, there is a need for a technique that can supply a data storage capability up to 4,000 bits with small weight, volume, and power requirements.

Recently, solid state storage devices have been used to provide small data requirements. The largest of those currently available is a 256-bit storage array that is made of p-channel MOSFET binaries. These binaries perform the same function as the magnetic core, that is, they store data and read it out upon command; however, they cannot perform counting or shifting functions. This array has the advantage of not requiring a large amount of peripheral drive circuitry in order to introduce and later retrieve data. It does, however, require some type of formating circuitry to present the data to the input of the memory and later collect it as it is retrieved. Furthermore, each bit dissipates 2 mW of power when it is being operated.

An alternate approach to solving this problem is to provide a shift register in which data may be stored until it is needed. GSFC has designed and fabricated a shift register, 128 binary digits (bits) long, on a single chip (Figure 1). Each shifting bit is made up of a master-slave flipflop that utilizes p-channel MOSFET switches and p-channel MOSFET resistors. The chip size is 125 mils by 128 mils, and it is presently being packaged in a 7-lead, round flat pack. It occupies a volume of 0.024 $in.^3$ and weighs 0.25 g. We have produced a number of experimental circuits which work as had been expected. The power dissipation is approximately 600 μW per bit,

or 70 mW for the entire register. The entire circuit is made up of more than 2,000 p-channel MOSFET's.

The two main applications for such a circuit are data storage and delay lines. A simple storage device could be constructed merely by putting a number of these circuits in either series or parallel configurations. Data may be shifted into a register when desired and retrieved at some later date. Data segments having various lengths may be accommodated merely by adding shift registers in series. In this way shift registers having varying lengths may be incorporated into a given storage array.

The circuit may be used as a delay line because a pulse presented at the input will appear at the output 128 clock pulses later. The duration of the delay from the input to the output may be changed by varying the frequency of the clock pulse. In addition to this, long delays may be accomplished by using low clocking speeds, by adding registers in series, or by combinations of both of these techniques.

This circuit, and circuits similar to it, can be produced in quantities up to 100. Should an application arise where greater quantities are required, the circuit can be manufactured by commercial semiconductor manufacturers. The masks have been designed to insure that they are compatible with the fabrication processes of at least two commercial producers. It is my estimate that this circuit could be mass-produced at a cost of $100 each, or less than $1 per bit.

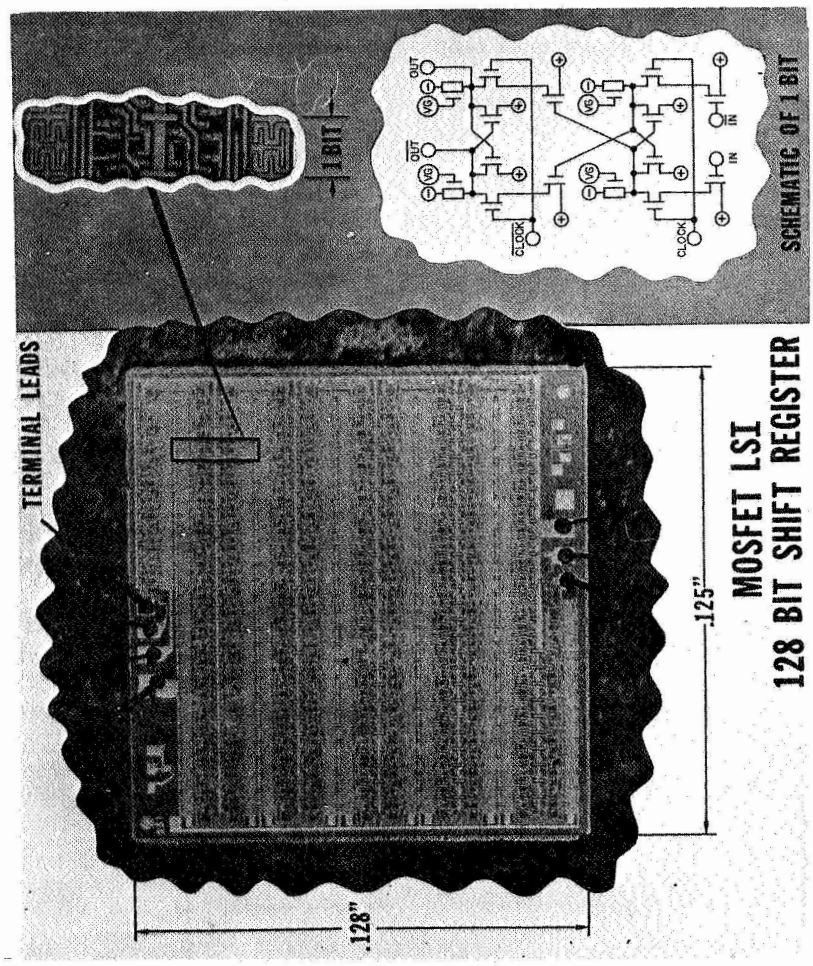

TERMINAL LEADS
1 BIT
SCHEMATIC OF 1 BIT
MOSFET LSI
128 BIT SHIFT REGISTER
.125"
.128"

Figure 1

DEVELOPMENT OF A SMALL COMPUTER FOR ON-BOARD PROCESSING

Rodger A. Cliff

CHAIRMAN:

Our next speaker is Mr. Rodger Cliff, and his topic is the development of a small computer for on-board processing.

MR. CLIFF:

Hard-wired, special-purpose data processors have been employed for many years on small scientific spacecraft. They are used, for instance, for data compression, data formating, and experiment control. As these hard-wired processors have become more complex, it has become more difficult to design and fabricate them.

We have been developing a stored-program computer called the SDP-3 with ART/SRT funds. This computer is of the same order of complexity as the previously employed hard-wired processors that it can replace. The SDP-3, however, has the advantage of flexibility that derives from using a general-purpose machine with a stored program. This one hardware design can be used in numerous applications by changing the program. It can even be time shared between several experiments.

We are fortunate to have an opportunity on the IMP I spacecraft to demonstrate the utility of the SDP-3. Figure 1 is a block diagram of the IMP I computer experiment. In this application the SDP-3 will handle data acquisition, buffering, compression, and formating for four of the scientific experiments. After launch, programs in the SDP-3 may be modified from the ground via the spacecraft's command system. Several of the experimenters plan to use this facility to optimize the operation of their experiments. It is not possible to have this degree of flexibility with hard-wired processors.

The SDP-3 is a serial machine which uses two's-complement arithmetic. It has a 4096 word memory which can be expanded to 65,536 words. There is one hardware index register and one level of indirect addressing. Communication with the experiments, the spacecraft command system, and the spacecraft data system is via serial input/output channels.

The SDP-3 will be programmed to use a variable format (shown in Figure 2) for transmitting the compressed experiment data. The advantage of this approach is that during periods of high activity the most important data

can be transmitted, whereas during periods of low activity—when there is less high priority data—other, less important data can fill in. The result is more efficient use of the telemetry channel.

The variable format employed by the SDP-3 in the IMP I computer experiment will be the first use of a data-dependent format for spacecraft telemetry. We are able to use this type of format for two reasons. First, we have in the SDP-3 the capability for generating such a format. Second, the IMP I telemetry system uses a long constraint length (48-bit) convolutional code, and therefore the bit error probabilities (10^{-6} to 10^{-8}) are for the first time sufficiently low to permit the use of this type of format.

We have just completed a redesign of the SDP-3. This redesign was undertaken after a study of the SDP-3 architecture revealed that significant savings could be made in both power drain and component count. For instance, the development of a novel multiplication algorithm and implementation of the necessary instructions in the SDP-3 permitted elimination of the MQ register, which is normally required for multiplication. Also, by converting parts of the SDP-3 from parallel to serial operation and then employing medium scale integration shift registers with eight stages per chip, the power dissipation and parts count have been further reduced. As a result of the redesign, the power dissipation of the SDP-3 (exclusive of the memory system) has been reduced by 30% to 1.3 W and the component count has been reduced by 30% to 500 integrated circuits.

We are looking forward to the launch of IMP I because it will be the first time that a stored-program computer has been used for this application. We are particularly enthusiastic about the facility for experiment control, post-launch reprogramming, and a data-dependent variable format.

THE IMP I COMPUTER EXPERIMENT

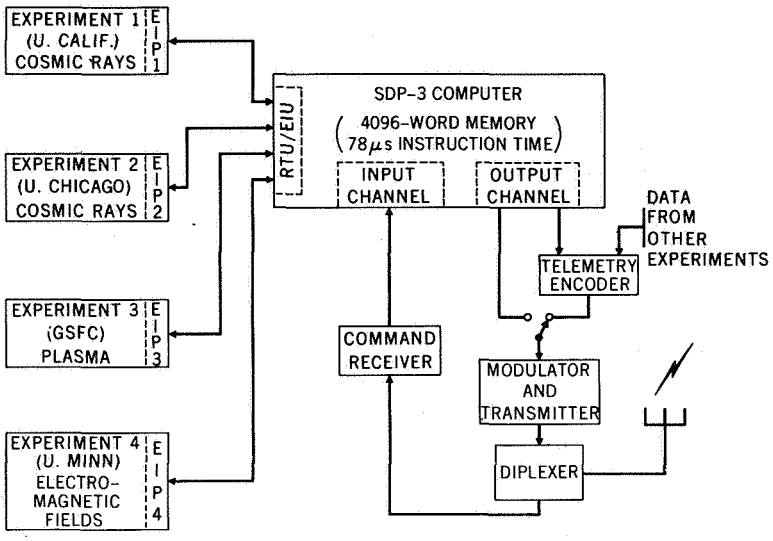
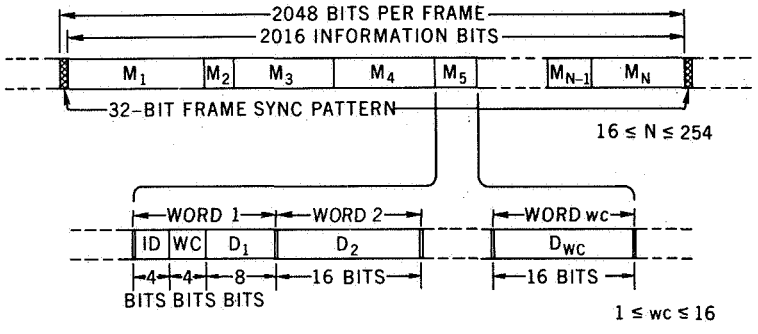


Figure 1

VARIABLE FORMAT FOR DATA TRANSMISSION



ID	MESSAGE TYPE
0	FILL
1	EXPERIMENT 1 (U. CALIF.)
2	" 2 (U. CHICAGO)
3	" 3 (GSFC)
4	" 4 (U. MINN.)
⋮	⋮
14	COMPUTER EXPERIMENT
15	REAL-TIME DECOM. AND PRINTOUT

Figure 2

THE BRAIN AS A MODEL FOR LSI

James S. Albus

CHAIRMAN:

The next paper will be presented by Mr. James Albus. His topic is the brain as a model for LSI.

MR. ALBUS:

Large scale integrated (LSI) circuits technology will soon make it feasible to consider building electronic systems containing many millions or even billions of active electronic components. Three questions then arise: What kind of systems should be built? What capabilities can be expected? What type of design techniques should be applied?

There already exists in nature a system that successfully and efficiently utilizes many billions of active components. This system is the brain.

I have been studying the component connection scheme of the brain and the data processing capabilities that result. One of the basic elements used by the brain appears to be a cell of the type shown in Figure 1a. If a pulse appears on the *in* line of this cell, the cell is activated and a pulse is sent out on the *out* line. If a pulse appears on the x input, the cell may or may not fire depending on the state of the x connection. If the x connection is in the 1 state, a pulse on *x* will trigger the cell; if the *x* connection is in the 0 state a pulse on *x* will not fire the cell. Consider the *x* connection to be normally in the 0 state and to be set to the 1 state by coincidence between a pulse on the *in* line and pulse on the x line. Thus, if an *x* line input is in the 1 state, it indicates that sometime in the past a coincidence occurred between an *in* pulse and an x pulse.

Now, visualize a cell with many variable strength *x* inputs as shown in Figure 1b. All the *x* inputs are independent; if a coincidence occurs between an x_2 pulse and an *in* pulse, only the x_2 connection will be changed. All the other connections will be unaffected.

It is possible to connect these types of cells into a planar array as shown in Figure 2. In this figure, each cell has its *in* line entering in the rear and its *out* line coming out the front. The x_1 lines from all the cells are connected to a single x_1 input; all the x_2 lines are connected to a single x_2 input, and so on.

If a pattern of impulses arrives on the *in* lines such that all the cells in the A-shaped darkened area fire, the "A" pattern will be transmitted via the *out* lines. If, while the "A" pattern is on the *in* lines, a pulse occurs on the X_1 line, the x_1 connections to the darkened cells will be set to the 1 state. Thus, at a later time, a pulse on the X_1 line will cause an "A" pattern to occur on the *out* lines even though nothing is coming in on the *in* lines. It can then be said that the "A" pattern is imprinted on the X_1 line. In a similar way a "B" pattern can be imprinted on the X_2 line, a "C" on the X_3 line, and so forth.

One possible use for this type of system might be in a trainable feedback control system. Consider, if the system shown in Figure 2 were placed between the pilot of an aircraft or spacecraft and the control surfaces or control jets, control commands from the pilot to the control mechanisms would appear as patterns on the *in* lines and be transmitted via the *out* lines. The X inputs would come from sensing devices such as airspeed indicators, inertial guidance sensors, and radar. Assume now the aircraft went into a dive. The inertial guidance would put out a train of pulses on X_1. The pilot sensing the dive would correct it by sending an "up" command to the control surfaces via the *in* lines. Coincidence between the "up" command and the X_1 pulses would imprint the "up" command on the X_1 dive sensor signal. Thus, at a later time, if the plane again went into a dive, the X_1 dive sensor signal would elicit an "up" command on the *out* lines even if the pilot were no longer present. In similar ways other maneuver commands could be imprinted on other feedback signals on the X lines.

This system, then, is a trainable feedback control system that can be trained by a pilot, possibly even under simulated conditions, and then substituted for a human pilot under flight conditions.

There are of course many unanswered questions of capacity, stability, and discrimination capability. However, it does appear that, by using extremely large numbers of components in the proper way, it is possible to develop electronic systems that exhibit brain-like capabilities.

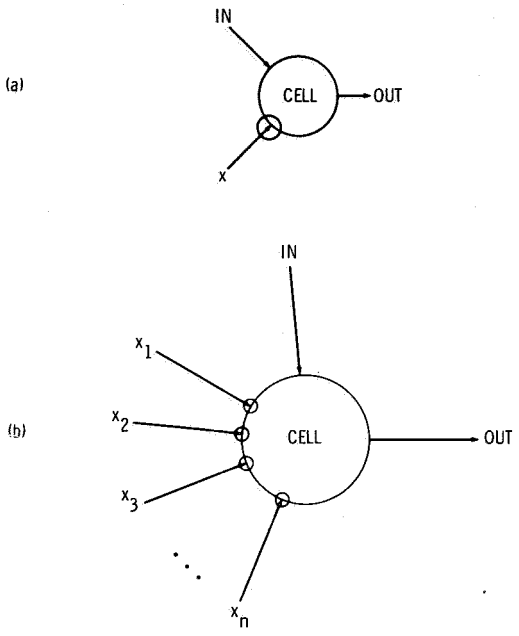


Figure 1

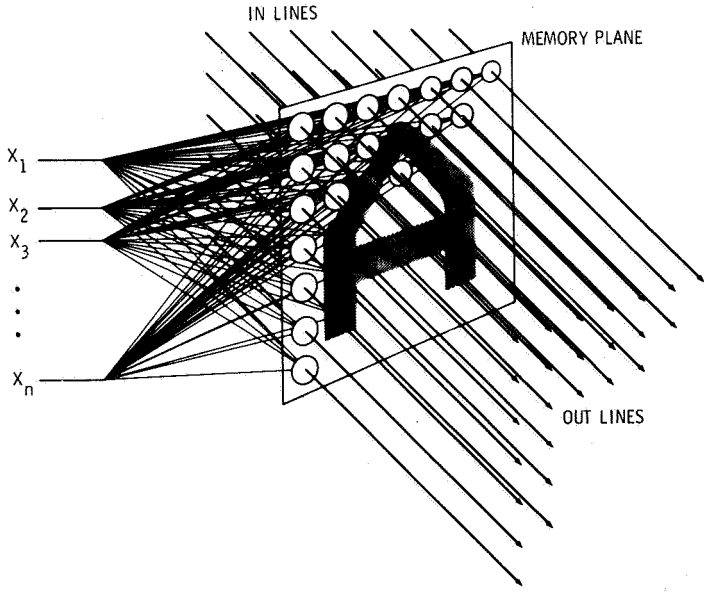


Figure 2

AN OPTICAL TELESCOPE FOR SPACE ASTRONOMY

John D. Mangus

CHAIRMAN:

Next, Mr. John Mangus will discuss an optical telescope for space astronomy.

MR. MANGUS:

The research program that I shall discuss today is directed toward the design, fabrication, and testing of efficient photon collectors for the ultraviolet (UV), extreme ultraviolet (EUV), and X-ray spectral regions. This past year has been devoted to research on and development of an optical telescope that can be applied to solar and stellar observations in the EUV region. The reason for concentrating our effort on EUV is illustrated in Figure 1.

Use of the classical Cassegrain system is limited to wavelengths greater than 1000 Å, the reason being that high reflectance coatings are not yet available for shorter wavelengths. To resolve this difficulty, we have implemented the glancing-incidence optical principle. Glancing-incidence optical systems, such as the X-ray (type I) telescope shown in Figure 1, have been built and used to obtain spectral as well as spatial information in the region from 1 Å to 100 Å.

On a geometric basis, the type I telescope appears to be useful as an efficient photon collector in the 100-Å to 900-Å spectral region. However, the narrow aperture that is an intrinsic characteristic of the type I telescope causes severe degradation of the optical transfer function, not only as a function of spatial frequency, but with increasing wavelength as well. For instance, the optical contrast function at 304 Å and at a typical maximum glancing angle of 1°is on the order of 40% to 60% for a type I telescope.

This difficulty has been resolved by implementing the optical system that was first proposed by Wolter for application in X-ray microscopy. In the past year, we have further developed its usefulness as an efficient collector in the EUV spectral region. The type II configuration, shown in Figure 1, consists of a parabola of revolution E_1 and a convex confocal hyperbola of revolution E_2. Since we are now reflecting longer wavelengths in the EUV, we are permitted to use larger glancing angles than those used with the type I telescope. Typically, the maximum glancing angle in the type II telescope will be of the order of 8° to 15°. Note that the line marked P in

each of the optical configurations is the principal surface, or nodal surface, for the telescope. Therefore, it is evident that an additional advantage of the type II telescope over the type I telescope is that its physical length may be shorter than its effective focal length.

Now, I will discuss some of the past year's developments in our research program. First, we have supplied the design for the EUV telescope in the GSFC EUV spectroheliograph instrument to be flown on OSO H. The spectroheliograph telescope, which is approximately 2 in. in diameter, has an effective focal length of 85 cm. Second, we have successfully fabricated this telescope in our laboratory. To my knowledge, it is the first time that this type of telescope has been made. Last September, a subcontractor also successfully fabricated this telescope. The reason for duplication of effort is that, at the initiation of the project, no one was quite certain that the telescope could successfully be made.

At the present time, test data on these telescopes indicate that they will have a resolving power of 5 to 10 seconds of arc at 304 Å. Geometrically, the resolving power of these telescopes is predicted to be between 3 and 4 seconds of arc at a half-field angle of 15 minutes of arc. Testing of these telescopes in the visible region has again indicated to us that additional work must be done on the diffraction theory of aberrations if we are to successfully predict the telescopes' ultimate performance. As an example, the off-axis image quality of the telescope does not degrade as rapidly as one would predict using geometrical theory, even when the wave abberration is of the order of several wavelengths. Other anomalies also have been noticed, such as the optical focal surface's depending upon the criteria used to evaluate the telescope. For instance, in the visible measurement, focal surface curvature is twice that predicted by geometrical analysis. The EUV data definitely indicate that the focal surface curvature determined from energy in a spot is half that predicted by a geometrical analysis. Measurements based on the resolving of two closely spaced targets indicate that the focal surface curvature is identical to that predicted by geometrical analysis.

Finally, the conditions for optimizing the image quality of the EUV, or type II, telescope have been established. A typical example is shown in Figure 2. Here, the telescope used has an effective focal length of 1 m and a collecting area of 15 cm^2. The graph in the upper left portion of Figure 2 shows that, if the telescope were pointed at the center of the Sun, the spot diameter would have to be of the order of 1.5 seconds of arc in order to enclose 100% of the radiant energy. Curve *B* in this chart is for the full width at half intensity. For this latter case, the spot diameter would have to be less than 1 second of arc.

In conclusion, the technology now exists to design and fabricate telescopes that will preserve spatial information as well as collect spectral information at wavelengths greater than 1 Å. Therefore, the program is now directed toward obtaining the developmental requirements of large-area (of the order of 500 to 1000 cm^2) collectors for solar and stellar applications in the EUV. Development of fast (small focal ratio) spectrometers for glancing incidence angles is also essential if the optimum properties of EUV telescopes are to be applied to nonphotographic missions. The physical diffraction theory of imaging must also be more thoroughly investigated if the results of photographic missions are to receive proper interpretation.

DR. PIEPER:

What are the requirements for surface finish and for the alignment of the optical parts in the type II telescope?

MR. MANGUS:

The surface finish should be as smooth as possible. When we design these telescopes, we require a specific polishing procedure that is called the bowl-feed technique. This technique has been evaluated by Bennett at China Lake. Typically, the rms roughness is of the order of 7 Å to 8 Å, if one uses a Cer-Vit material or the ultra-low-expansion fused silica that Corning makes. The alignment of the telescopes is critical. For the OSO experiment, we space the elements with a fused silica spacer. The secondary element is held by a beryllium spider. We try to keep the temperature variations on the spider less than 2° or 3° F.

Typically, we want the lateral displacement alignment of the two elements to be held within 2μ or 3μ.

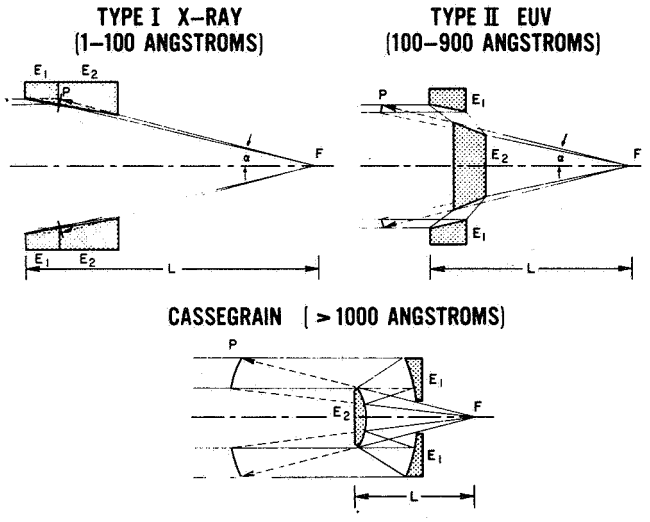


Figure 1

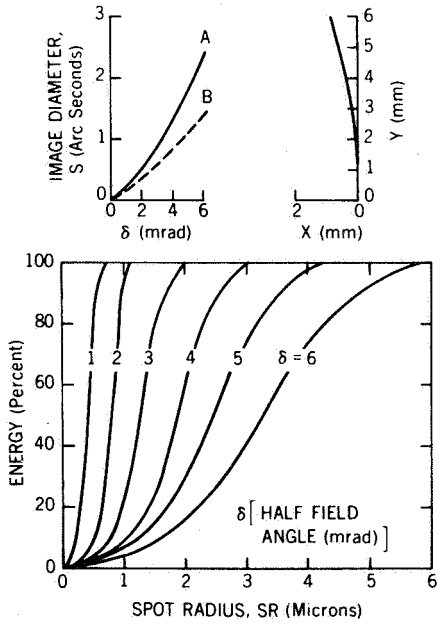


Figure 2–Performance data for a type II telescope with a collection angle of 15×10^{-4} sr and an effective focal length of 1 m.

IMAGE RECORDING DEVICES FOR ASTRONOMY

Kenneth L. Hallam

CHAIRMAN:

Dr. Kenneth Hallam will speak next about image recording devices for astronomy.

DR. HALLAM:

There is a need for efficient, high performance image recording devices for optical space astronomy. Such devices must meet certain performance requirements if they are to be generally useful. They must be compatible with conventional telescope configurations, such as those encountered at the focus of a Cassegrain telescope or a spectrograph. This puts real constraints on the physical configuration of such detectors. The linear resolution of the detector should be kept as high as possible in order to allow the overall format of the image on the detector to be minimized. This will provide a better limiting signal-to-noise ratio for each detector element and will at the same time permit more freedom in the telescope optical design, thereby improving its performance. At the same time, the total number of resolution elements of the detector should be sufficiently high to match the optical information content presented to it by the telescope or optical instrument. The detector must provide spectral selectivity, especially in the UV spectral region where suitable, long-wavelength, optical cutoff filters do not exist. In most photoelectric devices this selectivity can be provided by the choice of the photosensitive surface. Finally, the information output of the detector must be suitable for image data transfer and storage, either by photographic film or other registration media, or preferably by telemetry.

For the UV spectral region between 1050 Å and 3000 Å, we have successfully worked with industry over the past several years to produce a high-resolution UV image converter suitable either for direct photographic recording or for coupling to a television pickup. The device is shown in Figure 1. The optical image is focused onto the UV photocathode, which is supported on a thin plane LiF entrance window, transparent down to 1050 Å. Photoelectrons emitted by the cathode are accelerated by a 20-kV electric field, through a 200-G solenoidal magnetic field, to the anode at the other end of the converter tube, where they are focused onto an aluminized P11 phosphor. The phosphor makes the electronic image visible, producing an intensified blue-light image of the original UV scene. The blue image is

transferred through the fiber optic output window to the outside face of the tube with only small losses of light and resolution.

Initial development was concerned with establishing a suitable tube design, consistent with overall size and weight restrictions, that would yield high-resolution images of high uniformity and low distortion. Later development has extended the range of usable photocathode types, improved the permanent-magnet design, and provided for direct optical coupling of the output by means of fiber optic plates. The last is especially important, as otherwise a bulky and heavy optical relay lens must be used between the converter output and the photographic recording film or TV pickup, with the further consequence that at best only a small percentage of the light output would be transferred.

The performance characteristics of this magnetically focused UV image converter are shown in Table 1. Photocathode materials with long-wavelength cutoffs at 1500 Å, 1900 Å, 2200 Å, and 3000 Å are available. The front window of LiF transmits 60% of the light at 1216 Å. The quantum efficiency of the photocathode is typically about 10%, and the photon gain is about 60.

Table 1

UV Image Converter Characteristics

Characteristic	Value
Spectral ranges	1050 to 3000 Å
LiF window transmission (1216 Å)	60%
Window/cathode quantum efficiency	10%
Usable photon gain	60
Center limiting resolution	80 line-pair/mm
Edge limiting resolution	64 line-pair/mm
Opacity	10^4
Background (self noise)	$< 5\ \mu\mu L$
Input/output useful diameter	25 mm
Distortion	5%
Uniformity	±10%
Phosphor	P11 (4500 Å)
Voltage	25 kV
Magnetic field strength	210 G

The limiting resolution at a modulation transfer ratio of about 5% is between 64 and 80 line-pair/mm. These tubes have an opacity of 10^4 or better to directly transmitted light. The very small background noise produced by these tubes provides a favorable signal-to-noise ratio, which permits us to record faint objects.

The useful image field diameters of these tubes is 25 mm, but a new version can now be provided with a 40-mm diameter. The geometric distortion across the image is typically 5%, which is quite satisfactory for such devices. The uniformity of sensitivity across the image is of the order of 10% on good tubes.

These image converters are now ready to be used in rocket astronomy, where photographic recovery is feasible. Future work will be directed toward making them compatible with TV recording devices.

ULTRAVIOLET IMAGE CONVERTOR

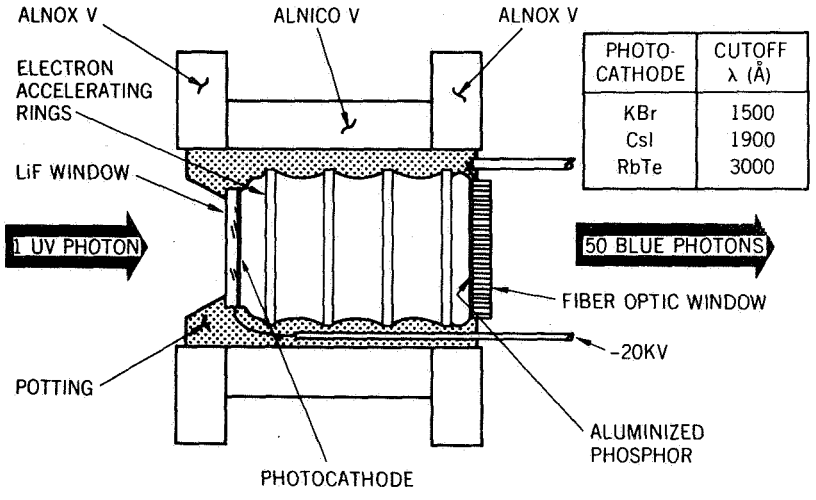


Figure 1

OPTICAL COATINGS FOR SPACE ASTRONOMY

John F. Osantowski

CHAIRMAN:

Our next speaker is Mr. John Osantowski, who will discuss optical coatings for space astronomy.

MR. OSANTOWSKI:

The development of high-efficiency, MgF_2- and LiF-protected Al reflectance coatings has contributed to recent advances in UV space astronomy. Unfortunately, this technology has not been implemented by industry to the extent that an individual experimenter can rely on commercial sources for specific requirements. Reflectance filters, transmission filters, and beamsplitters for use in the L_α region may be required for future applications. In addition, reflectance studies of vacuum-deposited thin films in the spectral region below 1000 Å are needed to improve the efficiency of mirrors and diffraction gratings.

We have programs to fill the gaps and requirements just mentioned, i.e., to provide the experimenter with state-of-the-art UV optical coatings and R&D for future applications. Three areas in which we have made significant accomplishments are—

1. Deposition of LiF-protected Al on the Princeton Experimental Package (PEP) prototype primary mirror.

2. Studies of UV reflectance filters.

3. Evaluation of MgF_2-protected (Al for use in the wavelength region below 1000 Å (Reference 1).

Figure 1 summarizes the results for the deposition of LiF-protected Al (Al + LiF) on the PEP prototype primary mirror. The radial reflectance distributions at L_α and L_β for a simulated mirror and for the PEP mirror are shown in the upper figure. Spectral reflectances for a typical Al + LiF sample and for a sample coated with the PEP mirror are compared in the lower figure. From test results on the simulated mirror, we conclude that the PEP primary can be coated with a reflectance uniformity of 2% or better at L_α and L_β. The results recorded for the prototype are lower due to outgassing from the mirror cell during the deposition, a problem we hope

to eliminate when the flight mirror is coated. Reflectance uniformity, however, is consistent with the test results.

Figure 2 illustrates the general characteristics of the UV reflectance filters we have studied. Construction of the filter is shown schematically on the right. Major features are the low-reflectance region in the visible spectrum and broad high-reflectance bands in the UV, i.e., the filter is solar blind. Reflectance peaks are located at 2600 Å and 1400 Å; maximum reflectances are 85% and 60%, and the widths at half-maximum are 1200 Å and 200 Å, respectively. The contrast (the ratio of peak reflectance to minimum visible reflectance) is 20 for the 2600 Å peak. If the filter is used in a double-pass mode, the reflectance is reduced, as shown by the lower curve. At 2600 Å, peak reflectance and bandwidth are reduced to 70% and 700 Å, respectively, and the contrast is increased to 400. To utilize the reflectance band at 1400 Å, a detector with low sensitivity in the near-UV region must be employed. A filter of this type was provided for the Nimbus 4 backscattered UV experiment to reduce scattered, visible radiation in the optical system.

Figure 3 demonstrates that MgF_2-protected Al, commonly used as a high-reflectance coating in the wavelength region above 1100 Å, can be modified to be useful at wavelengths down to 600 Å or 700 Å without completely sacrificing long-wavelength efficiency. The solid curve shows the reflectance for Al coated with 150 Å of MgF_2, which is about 100 Å less than that required for a reflectance of 80% or better at 1200 Å. For comparison, the reflectances of a Au film and of Ir films deposited at two substrate temperatures are given. These films are often used as high-reflectance coatings below 1000 Å. In the short-wavelength region, this comparison favors the protected Al sample over Au down to 600 Å and over Ir from 700 Å to 900 Å. Above 1100 Å the reflectance of the coated Al is about three times higher than those of the metallic films. However, due to the thinness of the protective layer of MgF_2, changes in reflectance as a function of age must be investigated before a final evaluation of this coating can be made.

REFERENCE

1. Hunter, W. R., Osantowski, J. F., and Hass, G., "Reflectance of Al Overcoated with MgF_2 or LiF from 1500 Å to 300 Å at Various Angles of Incidence," presented to the Optical Society of America, October 1969.

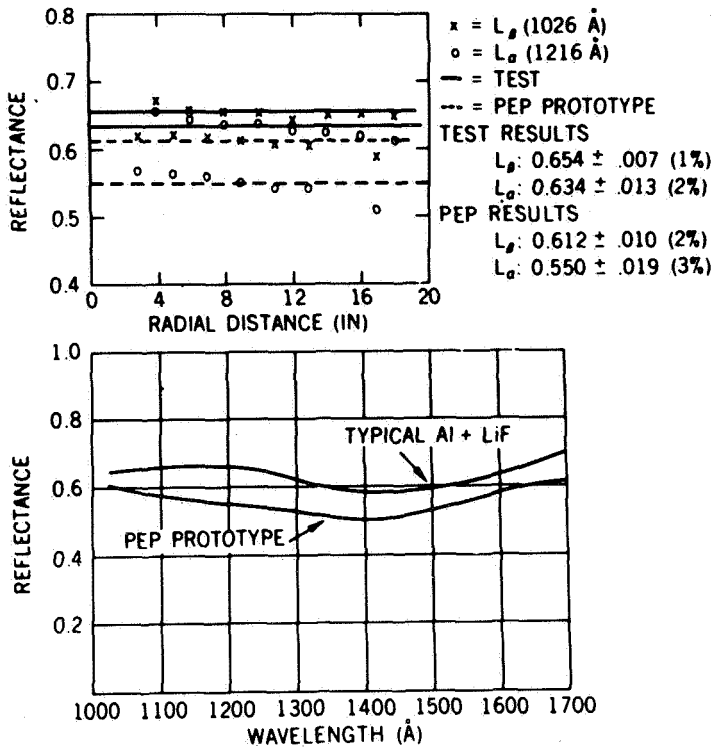


Figure 1–Results for Al + LiF deposited on PEP prototype primary.

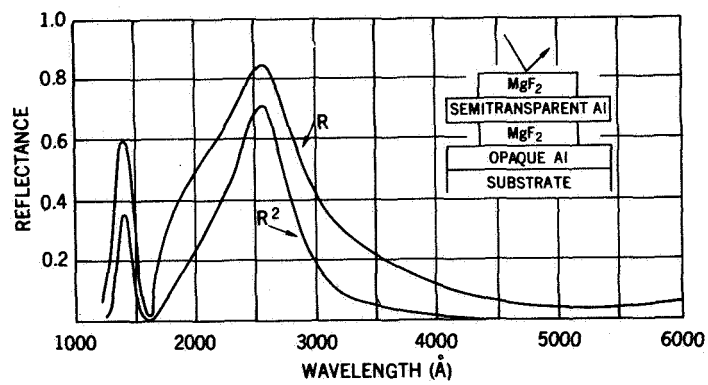


Figure 2–Normal incidence reflectance for MDM reflectance filter.

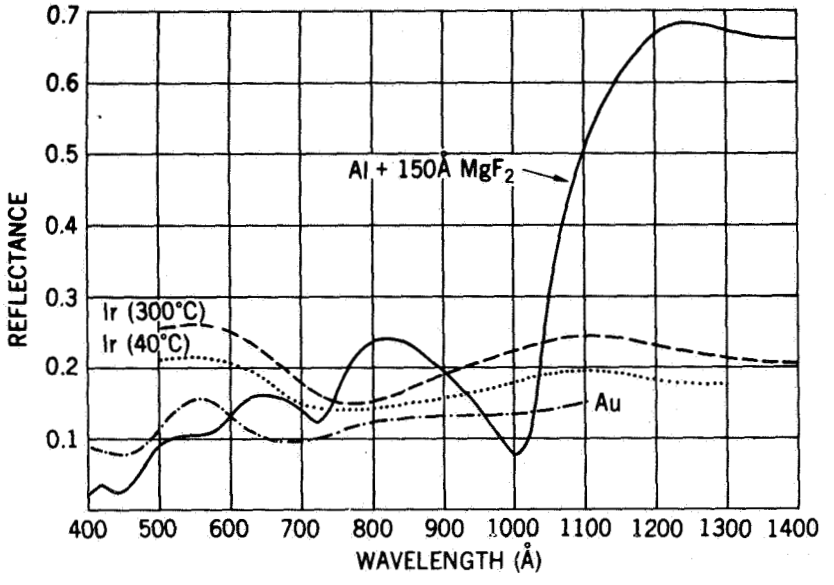


Figure 3– Comparison of the normal incidence reflectance for Al + MgF_2 (150-Å coating), Ir, and Au.

SOLID STATE DETECTOR DEVELOPMENTS

Donald J. Williams

CHAIRMAN:

Our next speaker, Dr. Donald Williams, will discuss solid state detector developments.

DR. WILLIAMS:

The charged particle detector has been and will continue to be one of the key instruments used in the pursuit of space research. One of the most reliable charged particle detectors flown aboard satellites and rockets in the last decade has been the surface barrier solid state detector. When used properly, this type of detector provides high-quality measurements over wide energy and flux ranges for extended periods of time.

For several years, we have conducted a continuing program designed to further improve and extend the particle detection capabilities of solid state detector experiments. Our main efforts have been directed toward the following areas of interest:

1. Low energy operation. We are particularly interested in covering the difficult region of from about 10 keV through a few hundred keV.

2. Multi-environment systems. Here, emphasis is on normal operation over a wide range of temperatures.

3. Long lifetime ($\geq$ 10 years) detection systems. Using information obtained from our in-flight and laboratory experience with these detectors, we can now construct experiments for 5- to 7-year missions if we test and store the detectors in our laboratory for at least 1 year.

4. Radiation damage effects. Our studies (References 1 and 2) in this area have uncovered ways of prolonging detector lifetimes in high radiation environments by four to six orders of magnitude.

5. Particle identification at low energies. Here, use is made of extremely thin [$\leq 5\mu$ ($\leq$ 200 μin.)] surface barrier detectors.

In the following discussion, I shall briefly describe our present solid state detector capabilities in areas 1 and 2.

Our attempt to obtain low-energy particle detection systems with solid state detectors has proceeded along two fruitful courses: use of surface barrier detectors and use of the avalanche diode. Surface barrier detectors offer the advantage of essentially windowless (negligible dead layer effects) operation, thereby permitting the efficient detection of electrons and heavy ions at low energies. The avalanche diode is a diffused junction device employing a bias field sufficient to cascade the free electrons generated by the incoming ionizing particles. A particle detector with internal gain is thus obtained. Gains of from 20 to 200 are currently available, yielding a favorable signal-to-noise ratio.

Figures 1, 2, and 3 illustrate our present capability for surface barrier detectors and avalanche diode devices. Figure 1 shows the spectra resulting from the incidence of 15-keV and 45-keV proton beams on a surface barrier detector 300μ thick and 25 mm^2 in area. The following items of interest apply to these results:

1. Energy thresholds of 15 keV can be routinely maintained with surface barrier detectors.

2. These results have been obtained at room temperature, i.e., nominal satellite operating temperatures. Cooling produces significantly improved results and lower energy operation.

3. Because of dead layer effects, the low-energy operation shown for protons in Figure 1 cannot be attained by diffused junction detectors.

Figure 2 shows the electron detection efficiency of the same surface barrier detector. This efficiency rises rapidly from ~ 10 keV to a nearly constant value above ~ 20 keV that is determined primarily by backscattering of the incident electrons out of the detector. The surface barrier detector results shown in Figures 1 and 2 would not have been possible without a continuing parallel effort in preamplifier development in our laboratory. Flight packaged preamplifiers are now available having a room temperature noise of ~ 1.3 keV (full width at half maximum) with 0 pF input capacity and a noise slope of ~ 13 eV/pF.

Figure 3 shows the electron efficiency of an avalanche diode device. It can be seen that this curve is quite similar to the surface barrier detector curve. However, the avalanche diode has two main disadvantages when compared to the surface barrier unit. The avalanche diode has a dead layer problem and cannot detect protons with energies $\lesssim$ 100 to 150 keV, and the avalanche diode has very poor energy-resolution characteristics.

However, the avalanche diode has the useful property that its low-energy particle detection characteristics are insensitive to the ambient temperature. The efficiency curve shown in Figure 3 can be duplicated at 60° C, and we are preparing to perform tests at 80° C to 100° C. Thus we have a device that is potentially capable of performing efficient low-energy electron and X-ray measurements at high ambient temperatures and that requires no special spacecraft cooling techniques on close-in solar fly by missions.

In summary, low-energy (~ 15 keV) charged particle detection is possible using surface barrier detectors, and the avalanche diode shows great promise as an efficient low-energy device capable of operation over a wide range of temperatures. Our program continues to strive for similar increases in detector capability for all the areas of interest mentioned previously.

MEMBER OF THE AUDIENCE:

Can you say something about energy resolution?

DR. WILLIAMS:

The avalanche diode at present has no energy resolution. It is operated in a saturated mode. The problem of energy resolution in such a device is to maintain a uniform multiplication region, which is extremely difficult to do. It relates to the purity of the silicon, and one of the things we have found is that it is very hard to convince the growers of silicon to maintain very tight controls on the quality of the boules that they grow.

REFERENCES

1. Coleman, J. A., Love, D. P., Trainor, J. H., and Williams, D. J., "Low Energy Proton Damage Effects in Silicon Surface Barrier Detectors," *IEEE Trans. Nucl. Sci.* 15 (1): 482-490, 1968.

2. Coleman, J. A., Love, D. P., Trainor, J. H., and Williams, D. J., "Effects of Damage by 0.8 MeV-5.0 MeV Protons in Silicon Surface Barrier Detectors," *IEEE Trans. Nucl. Sci.* 15 (3): 363-372, 1968.

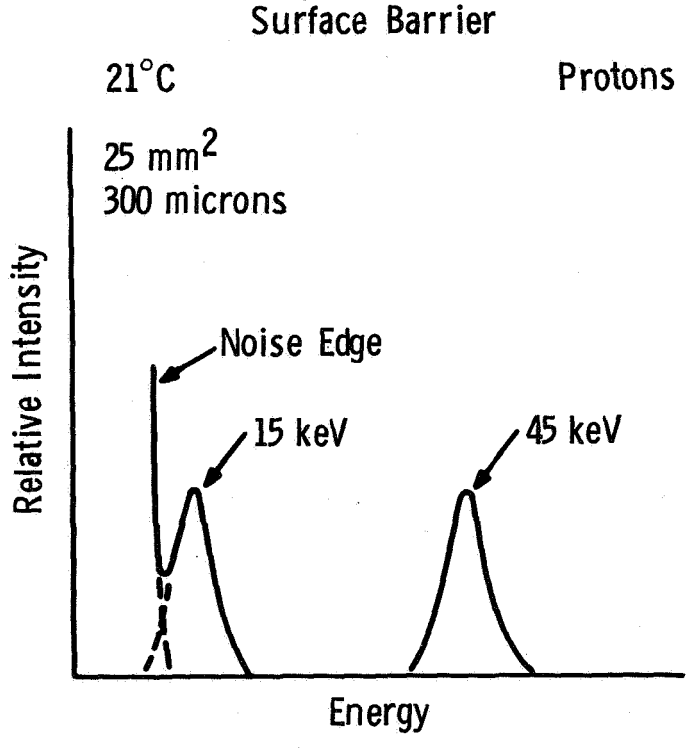


Figure 1

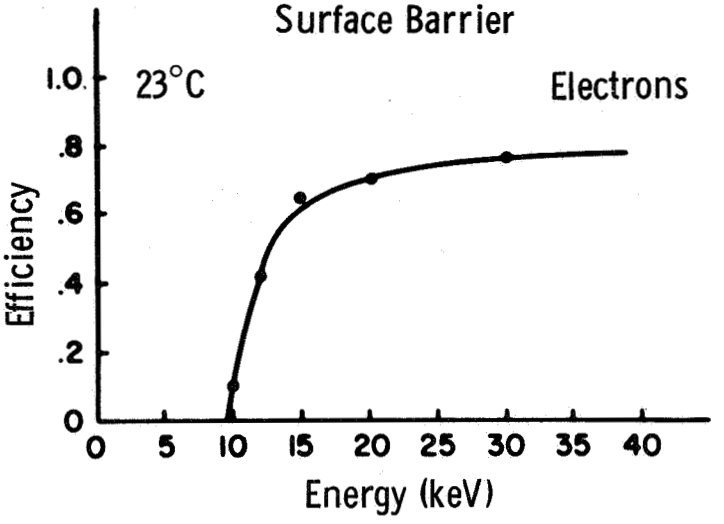


Figure 2

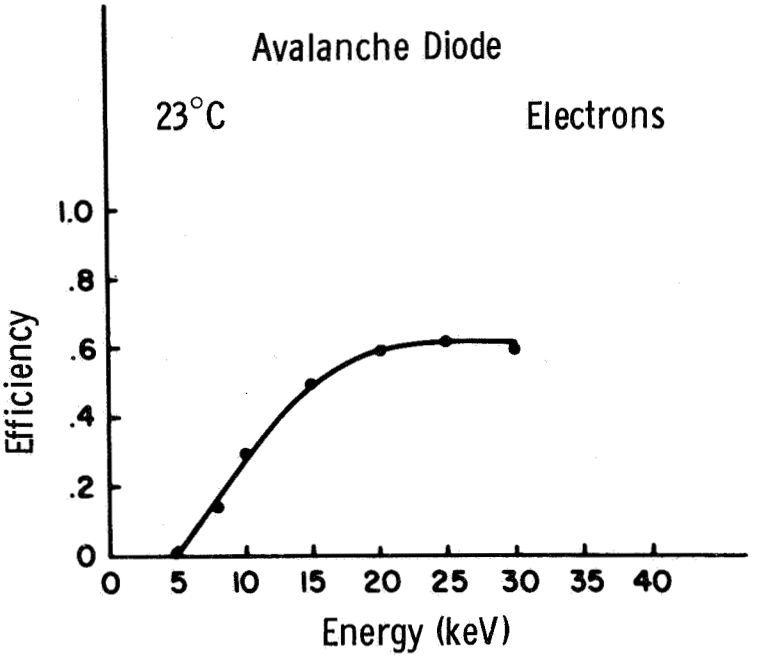


Figure 3

THE INTERROGATION, RECORDING, AND LOCATION SYSTEM EXPERIMENTAL RESULTS

Charles E. Cote

CHAIRMAN:

The next area to be discussed is spacecraft systems. Mr. Charles Cote will review IRLS tracking results.

MR. COTE:

The interrogation, recording, and location system (IRLS) is a satellite technological experiment designed to demonstrate the feasibility of position determination and scientific data collection from remote instrumented platforms deployed on a global scale.

The system is depicted in Figure 1. Addresses of platforms expected in view are programmed into the satellite on an orbit-by-orbit basis from a central ground acquisition and command station. Interrogation commands are executed throughout the orbit at predetermined intercept times. Platform locations are obtained by range-range measurements which yield solutions based upon intersections of mathematically generated ranging spheres with the surface of the Earth. Up to 28 channels of sensory data are collected simultaneously from each platform.

The experiment commenced with the launching of the Nimbus 3 satellite in April 1969 and has been in daily operation ever since. Experiments have been conducted with a total of 10 platforms located on balloons, ships, ice islands, aircraft, buoys, and land. The various platforms have been deployed on a global scale, encompassing the North and South Polar regions, the continental United States, Puerto Rico, Bermuda, and the Pacific Ocean. The installation and operation of the platforms have been under the control of the participating experimenters, and GSFC has been responsible for satellite operations and ground processing of received data. Since launch, over 3×10^6 bits of engineering, meteorological, oceanographic, and other scientific data have been collected from platform sensory devices; data quality has been excellent throughout.

Analysis of ranging and location data has shown that the ranging technique has performed according to design specifications. However, the quality of satellite-orbited ephemeris data has a marked impact on platform location accuracy. Typical location error trends are shown in Figure 2 for two stationary platforms. The plot represents orbit-by-orbit location errors

vs time from ephemeris update (designated epoch) for a 14-day period. As shown, the errors increase at a rate of approximately 1.5 n. mi. per day. Upon inserting updated orbital elements, location errors are substantially reduced, as shown in Figure 3. Under conditions of accurate ephemeris data, the system design goal of ±1.5 km has been demonstrated.

An example of location data obtained for moving platforms is shown in Figure 4. The plot represents movements of a drifting ice island in the North Polar region during the month of May. For comparisons, the solid line represents island positions as determined by the Navy Transit Satellite and relayed in IRLS data frames. Corresponding IRLS positions taken near epoch times are shown by *X*'s. The differences are of the order of 1 to 2 n. mi. In addition, each collected data frame contained eight readings of seismometer recordings as monitored by the platform.

Similar results were obtained during tracking operations using ships at sea and balloon flights across the United States. Experiments with higher speed aircraft (300 kt) have accuracies of ±10 n. mi.

An additional but unanticipated application of IRLS occurred in August 1969 when the Naval Oceanographic Office's Puerto Rico buoy was inadvertently separated from its anchor. Although the buoy was equipped with a shore radio alarm system, the device failed to operate and the buoy began drifting freely. With the aid of the IRLS, the buoy was easily located and recovered (Figure 5). Had IRLS not been aboard, it was estimated that the buoy would have been permanently lost. It is interesting to note that the IRLS system detected the presence of swirls and eddies off the southwest tip of the island during a 5-day period.

Experiments with free drifting buoys are planned for upcoming months. In addition, plans are under way to instrument an elk in Yellowstone Park to show the feasibility of animal tracking via satellite. Meanwhile, performance analysis will continue.

INTERROGATION, RECORDING, AND LOCATION SYSTEM (IRLS)

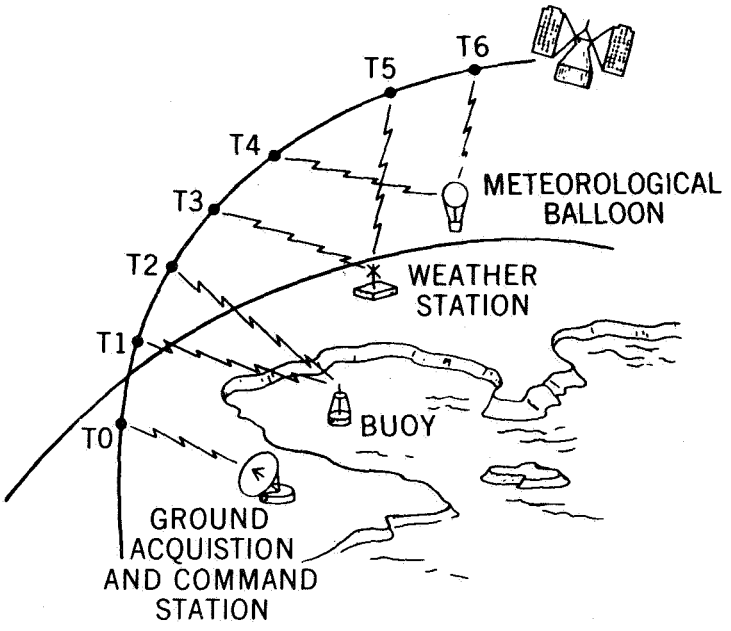
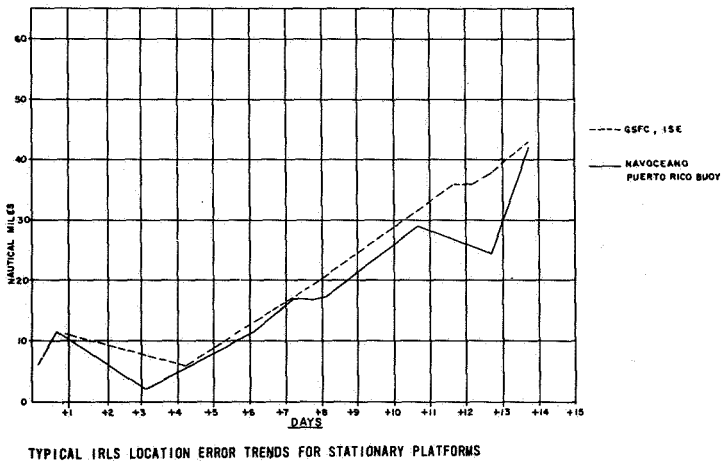


Figure 1



TYPICAL IRLS LOCATION ERROR TRENDS FOR STATIONARY PLATFORMS

Figure 2

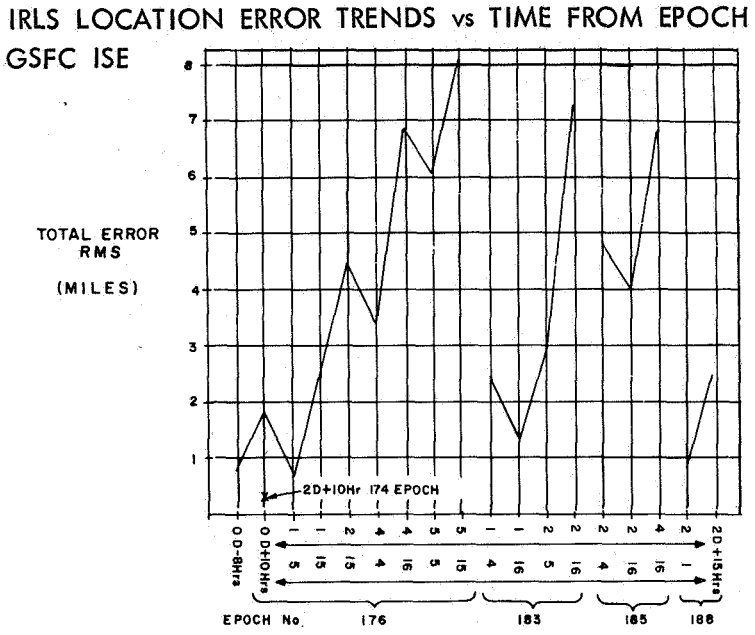


Figure 3

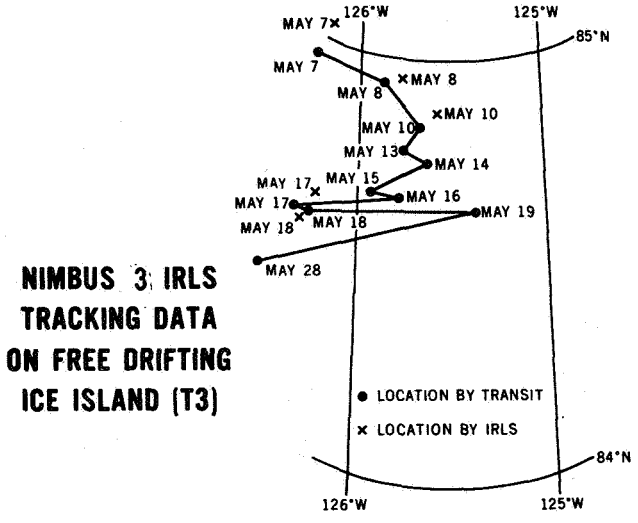
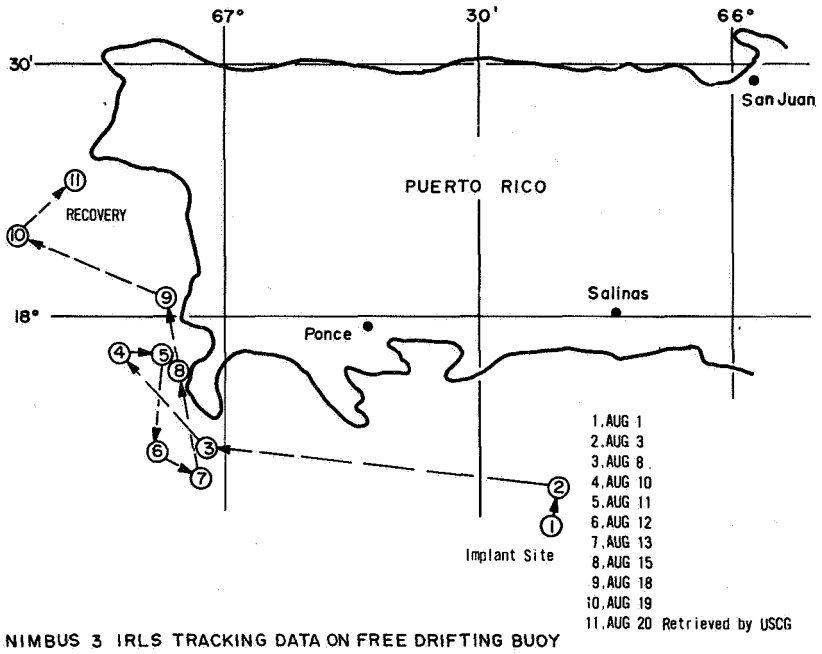


Figure 4



NIMBUS 3 IRLS TRACKING DATA ON FREE DRIFTING BUOY

Figure 5

A TROPICAL WIND TRACKING TECHNIQUE

Albert E. Arndt

CHAIRMAN:

Next, Mr. Albert Arndt will discuss a technique for tracking tropical winds.

MR. ARNDT:

Meteorologists indicate that the region of the tropics is most important to our everyday weather and requires additional exploration. For this reason a tropical wind experiment was initiated to study the monitoring and mapping of the wind fields in the tropics at three altitudes. Data would be randomly transmitted to a satellite from a global equatorial system of free-floating balloons. The data obtained in this program can be essential ingredients in the Global Atmospheric Research Program (GARP).

Each of the free-floating balloons would carry an electronic platform package capable of transmitting an identification signal, temperature and pressure signals, and a carrier signal for tracking the balloon. As the satellite passed through the tropical region (Figure 1), the data transmitted from the platform would be received, stored, and transmitted later when the satellite passed over a ground station. The collected data would be sent to GSFC for preliminary processing and then to the National Center for Atmospheric Research (NCAR) for detailed analysis and wind-field mapping.

An in-house study of applicable system concepts was conducted. Because of the large number of balloon platforms anticipated, cost was established as the primary constraint. Therefore, it was decided that only the essential functions would be performed on the platforms. Two approaches were considered in detail: a system in which the platform transmissions would be controlled from a satellite through an RF link, and a system that would transmit without any control from a satellite. From the standpoint of cost, the latter approach appeared to be the most attractive. Such a technique would provide a one-way link from each platform to the satellite.

From system definition studies, it was concluded that, for a system with sufficient capacity, each platform would be turned on for about 0.5 seconds, during which time the data would be transmitted, and then turned off for a period of about 50 seconds (Figure 2 illustrates the simple case of three transmitters). To facilitate the multiple access, that is, the handling

of signals from many transmitters to a single receiving source, each platform's transmission would be randomly spaced in time and frequency. This would be accomplished by using a low-precision timer that keys the platform on and off. In addition, by using a low-precision platform oscillator, the transmit frequency would be randomly spread about a central frequency. In this way, signals received at the satellite would be distributed randomly in frequency and emission time.

Tracking of each balloon by the direct determination of range is not possible with an incoherent, one-way transmission link. However, by measuring the frequency of the received carrier over an interval of time and noting changes in frequency from successive measurements, tracking information can be obtained. This change in frequency would be directly proportional to a change in range between the platform and the satellite.

The probability of failure (i.e., interference between transmissions) has been evaluated by simulations and plotted in Figure 3. Failure, in this case, is defined as not receiving at the satellite at least three good transmissions during a pass. It was found from our studies that a maximum of 10 transmissions is possible for those balloons located near the satellite subtrack and a minimum of five for those balloons located at the horizon. The plot shows that, for those balloon platforms located where 10 transmissions are possible, the probability of failure is insignificant, and it is still very low for those located where only five transmissions are possible. For the case shown in Figure 3, the time of transmission was allowed to vary randomly up to 10 seconds, and the limit of the frequency variation was ±15 kHz.

In conclusion, studies have shown that it is feasible and practical to utilize an incoherent, random transmission technique for receiving data and tracking a large number of balloon platforms.

TROPICAL WIND MEASUREMENTS BY BALLOONS AND SATELLITE

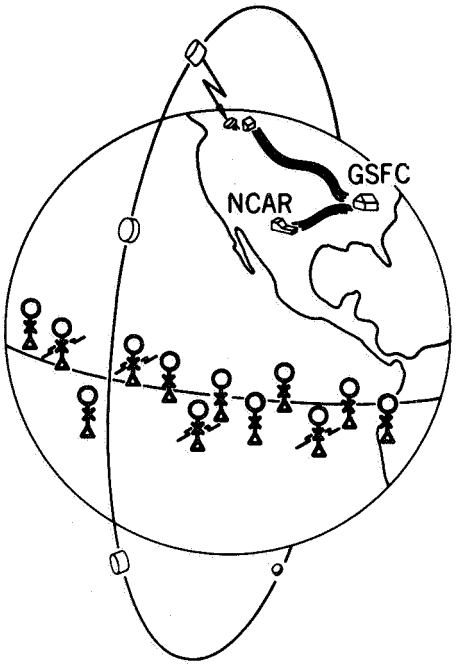


Figure 1

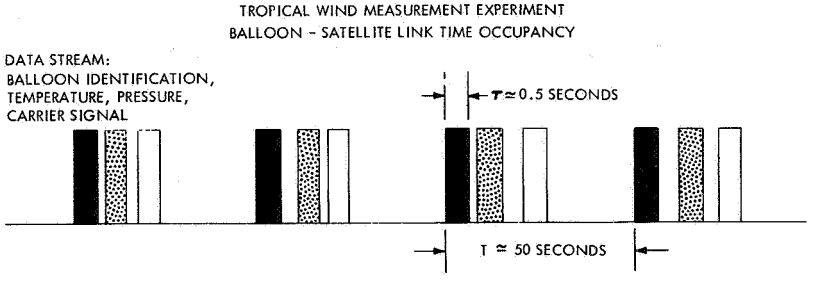


Figure 2

TROPICAL WIND MEASUREMENTS EXPERIMENT

PERFORMANCE CHARACTERISTICS

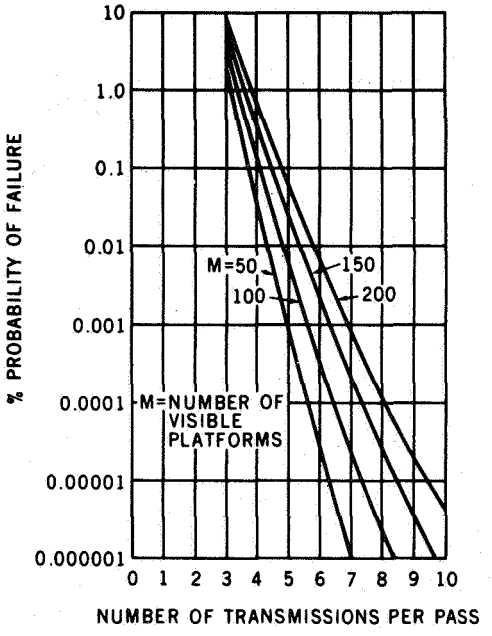


Figure 3

BROADENING OF SPECTRAL LINES IN A SOLAR SIMULATOR HIGH-PRESSURE ARC

Mathew P. Thekaekara

CHAIRMAN:

The next two papers have to do with the subject of spacecraft testing. Dr. Mathew Thekaekara will present the first, on the broadening of spectral lines in the high-pressure arc of a solar simulator facility.

DR. THEKAEKARA:

High-pressure mercury-xenon arc lamps are used as sources for radiant energy in GSFC's space environment simulator. The arc spectrum typically shows a certain number of strong lines superimposed on a continuum. The profiles of these lines afford a very helpful clue to an understanding of the nature of the mercury-xenon plasma and of the discharge mechanisms in these lamps, which offer convenient and effective means of solar simulation. The spectrum of mercury, unlike that of most elements, consists of a few widely spaced lines, and this property often permits spectral profiles to be examined without line interference. The lines occur mostly in the visible and near-infrared regions of the spectrum, thus making the arc effective as a solar simulator. An experimental program was undertaken to study these line profiles, and several varied and complex phenomena were observed. It was seen that these phenomena can be explained from the quantum mechanical theory of the mercury atom and from the statistical theory of the plasma.

Figure 1 shows four typical line profiles. The source is a mercury-xenon lamp, operating at 2500 W, with internal pressure about 30 atm. The marks on the 5770 Å and 5791 Å lines show where the unbroadened line from a low-pressure discharge source, for example a pen-ray lamp, would be. All lines are asymmetrically broadened, but no two lines are exactly the same in shift or width. The shift is to the red for most lines. For two of the lines shown in the figure, 5770 Å and 5791 Å, and for another pair at 3130 Å from the same upper levels (not shown), the shift is to the violet. The self-reversal dips also show asymmetric broadening and shift. Quantitative studies were made on about 20 lines at five power levels varying between 500 and 2800 W. The effects show a strong dependence on the input power and on the upper and lower levels involved in the transitions.

Figure 2, which shows the theoretically predicted line profiles, helps to relate these various profiles to the discharge parameters of the mercury-xenon arc. The narrow curve shows the line produced by a low-pressure

lamp. The width is mainly due to the finite resolution of the detector instrument. Line broadening is distinguished in three main types: Lorentz broadening, Doppler broadening, and quadratic Stark effect broadening. Lorentz broadening, which is due to the interruption of a wave train by collision, or Doppler broadening, which is due to the velocity of the atom along the line of sight, both yield a profile exactly symmetrical with respect to the unbroadened line and practically identical for all the lines in the spectrum. Quadratic Stark effect broadening is due to the emitting atom being in an electric field, such as that of ambient ions. It can be explained from the second-order perturbation theory of the quantum mechanical treatment of atomic spectra. The external field causes the energy levels to be shifted, the upper levels being more shifted than the lower. The effect due to a static field on the line produced by a single atom is to shift its position. The shift is proportional to the square of the field. Since a large number of transitions, nearly 10^{20}, occur per second, the line profile is the combined effect of many shifted lines.

The statistical probability of an atom being subject to an ionic field $\mathbf{F}$ varies with the value of $\mathbf{F}$. The distribution function of the probability $P(|\mathbf{F}|^2)$ has been derived by Holtsmark and is known as the Holtsmark distribution. If all the atoms and ions were stationary, this would also be the line profile, asymmetrically broadened and wholly to one side of the unperturbed line. Actually the field changes rapidly; energy level change causes phase change in the wave; the effect is that collision profile is superimposed on each element of the Holtsmark profile. A further modification is the self-absorption due to the relatively cooler vapor near the envelope of the bulb. Quadratic Stark effect occurs also here, though to a smaller extent.

The difference in line shift and width between lines is explained by the equations

$$\Delta E_s^0 = \frac{e^2}{c^2} \sum_{m \neq s} \frac{|L_{ms}|^2}{E_m^0 - E_s^0} |\mathbf{F}|^2,$$

and

$$L_{ms} = \int_\tau \psi_s^{0*} R\cos\theta \, \psi_m^0 \, d\tau .$$

The change in energy E_s^0 of a level s is due to all other levels which can optically combine with it, that is, all levels for which the matrix element L_{ms} of dipole transition is nonzero. The shift is up or down depending upon whether the main perturbing levels of energy E_m^0 are below or above the

level s. This explains why the shift is to the red in most cases but to the violet for two pairs of lines. The close correlation of the line profile to the upper and lower levels involved in the transition is explained by the equation.

It would seem that the spectral lines of the high-pressure mercury-xenon arc show some of the most interesting examples of line broadening and line shift available to a spectroscopist. The effects are many times stronger than in ordinary atomic spectra. They give a striking experimental confirmation of several aspects of the theory and yield information on the temperature, ion density, and other parameters of the mercury-xenon plasma.

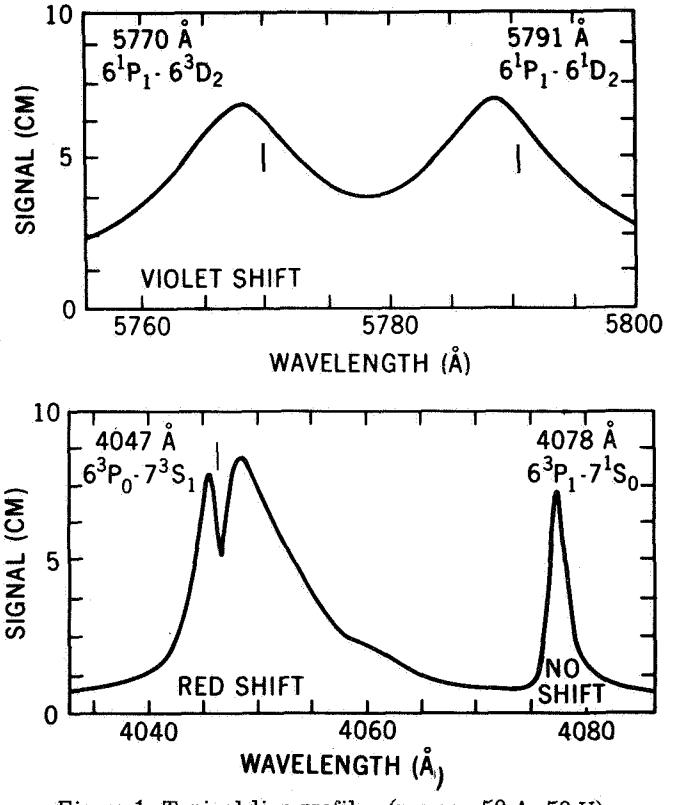


Figure 1–Typical line profiles (power: 50 A, 50 V).

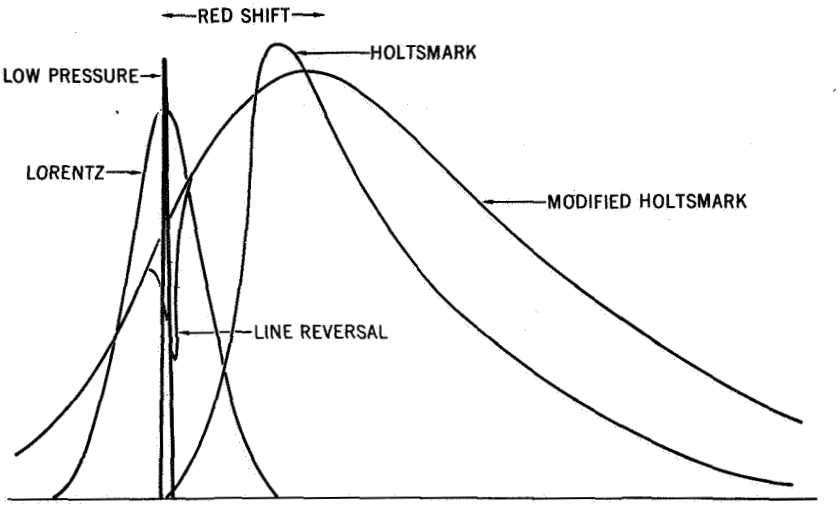


Figure 2–Theoretically predicted line profiles.

A TECHNIQUE FOR THE DETERMINATION OF THE DIPOLE MOMENT OF A SPACECRAFT FROM A LIMITED AMOUNT OF NEAR-FIELD DATA

William L. Eichhorn

CHAIRMAN:

The next paper will be by Mr. William Eichhorn, and he will discuss a near-field technique for the determination of the magnetic dipole moment of a spacecraft.

MR. EICHHORN:

A knowledge of the magnetic dipole moment of a spacecraft is necessary to determine the torque it will experience in orbit due to the Earth's magnetic field.

In the past this moment has been calculated from measured values of the spacecraft's magnetic field at distances that were large compared with the physical size of the spacecraft itself. At these distances, the spacecraft's field is purely that of its dipole moment, and the calculations are easily done.

However, as can be seen from the far-field data in Figure 1, the field level seen at these distances is fairly weak, and sometimes difficult to measure. This, together with the instrumentation noise (also shown in the figure) and the limited resolution of the measuring instruments, results in large uncertainties in the calculations. A further complication is that, in some cases, the spacecraft is so large that the distances required for a far-field measurement lie well beyond the boundary of a controlled magnetic environment and therefore cannot be attained.

In the cases described, the data must therefore be taken close to the object. In this region, the magnetic field, while being well increased in magnitude and relatively noise free, is also non-dipolar in nature. Therefore the dipole equations used in the far-field calculations cannot be applied, and other techniques are required to analyze this type of data. Several methods are available to do this, but the data required are difficult and tedious to obtain.

In every spacecraft test of this type, the only data points readily accessible lie in the horizontal plane. One can easily measure the spacecraft's field anywhere in this plane by simple rotation of the spacecraft about its vertical axis. Figure 2 shows such a data collection system. To

overcome the problems present in existing methods, we sought a specific technique for analyzing this type of data. It was found that if we could make an assumption about the nature of the spacecraft's magnetic field, a two-step procedure could be used to determine the dipole moment to very good accuracy. This procedure is diagramed in Figure 3.

The data from each of the magnetometer sensors are first Fourier-analyzed to obtain the first three Fourier coefficients. These coefficients are then used to solve a set of simultaneous equations for each of the three dipole moment components.

Extensive testing on samples whose magnetic dipole moments were known showed that this technique could give answers correct to within about 6 percent. When far-field analysis techniques were applied to the same data, the errors were of the order of 150 percent.

This procedure was applied during the magnetic tests of the ATS 5 spacecraft in July 1969. It allowed the spacecraft's moment of 5900 pole-cm to be compensated and reduced to 366 pole-cm.

The numerous calculations for this test were done on a desk calculator and took approximately 45 minutes each. A computerized version of this program presently being prepared is expected to greatly reduce the time needed to determine the magnetic state of a spacecraft.

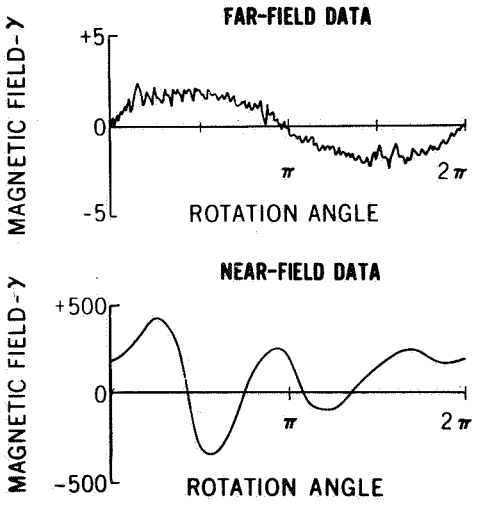


Figure 1—Far-field data compared with near-field data.

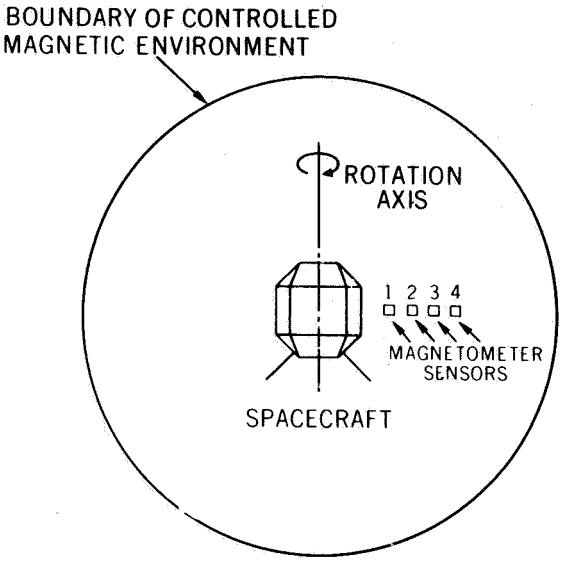


Figure 2–Near-field data collection system.

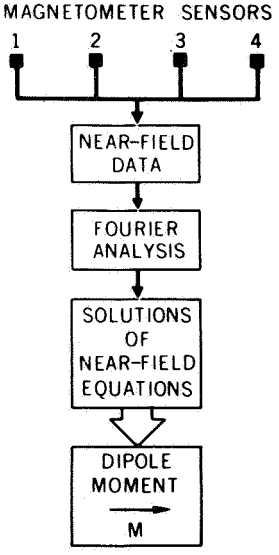


Figure 3–Near-field analysis procedure.

PULSE LASER RANGING DEVELOPMENT

Thomas S. Johnson

CHAIRMAN:

The next series of papers deals with ground systems, such as tracking telemetry, and timing. Mr. Thomas Johnson will speak on the development of pulse laser ranging techniques.

MR. JOHNSON:

During the past several years, we have been developing techniques and instrumentation to improve the accuracy, reliability, and operational efficiency of our laser ranging system. An improved modular ruby laser head was developed last year, and a great effort has been directed toward the investigation and improvement of the receiver-detector system.

An interesting application of these techniques occurred in the development of the acquisition ranging system for the Apollo 11 Lunar Laser Ranging Experiment. The astronauts left behind an array of 100 corner reflectors to serve as a fixed fiducial point on the lunar surface (Figure 1). A laser ranging system (Figure 2) capable of a 10-J pulse every 3 seconds was developed for use with the 120-in. telescope at the Lick Observatory. The system included beam correcting and coupling optics, boresight capability, a special receiver-detector assembly, and a ranging and data control system.

The Lunar Laser Ranging Experiment, compared with satellite ranging, presented a somewhat unique detection problem. The signal received from the reflector array was expected to be no larger than one to five photoelectrons, whereas the solar-illuminated lunar surface presented a very high noise background pulse rate. In order to discriminate against this high noise background and yet maintain high sensitivity and range accuracy, enhanced photomultipliers operating in coincidence were used (Figure 3). Since the noise background pulses occur randomly in the two tubes, only a very small percentage of the pulses would occur in coincidence, but the tubes would always have a simultaneous output pulse caused by the laser signal. The ranging measurement was made with a 1-nsec-resolution time interval unit started by a sample of the laser output and stopped by the coincident signal output.

We operated the system on the last night (August 3, 1969) that the 120-in. telescope was available for the experiment. During a period of one and

one-half hours, 1200 shots gave 100 apparent measured ranges. The data were analyzed by subtracting the measurements from the predicted range ephemeris and then correcting these residuals for parallax and other errors existing in the ephemeris.

The correction was done by fitting a first-order equation to the range residuals and recognizing that the true ranges would occur within the system precision with respect to the corrected line. The corrected data were then displayed in a histogram (Figure 4) with 100-nsec intervals over the range –5 to +5 μsec with respect to the corrected line.

The probability of having many noise pulses occurring within the single 100-nsec central interval is very low. It can be concluded that the pulses in the central interval are indeed from the reflector. The reflector range measurements exhibited an rms of 35 nsec, or about 5 m, which remains the most accurate actual measurement to the reflector package.

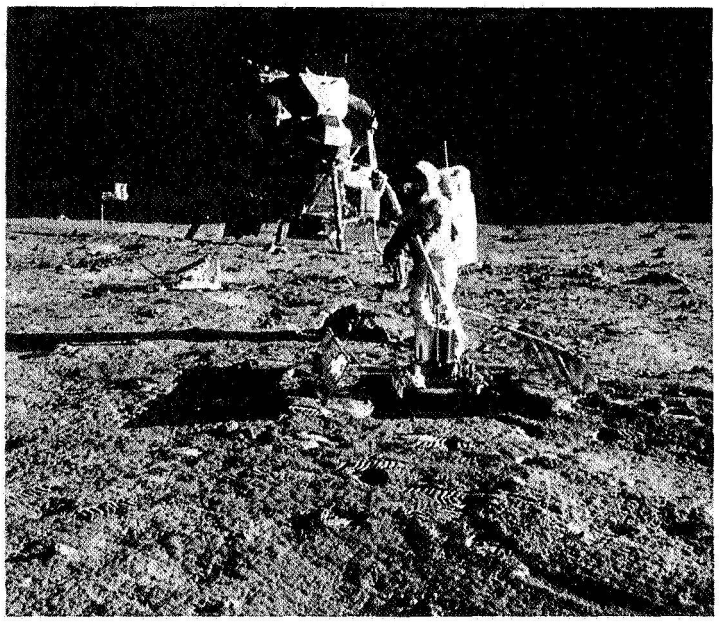
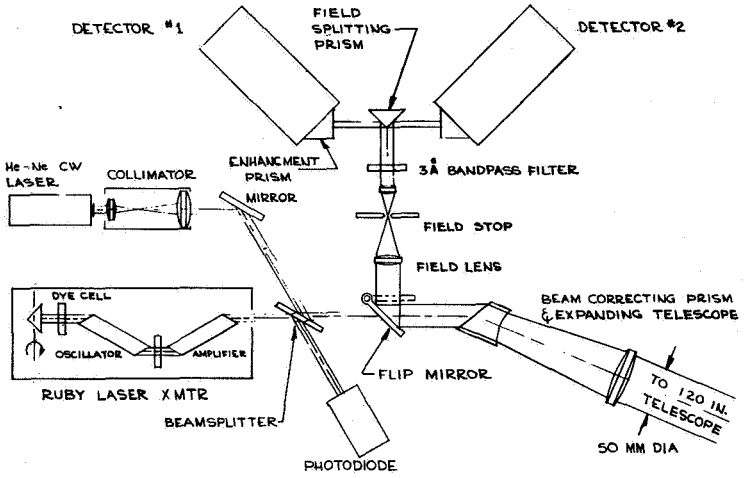
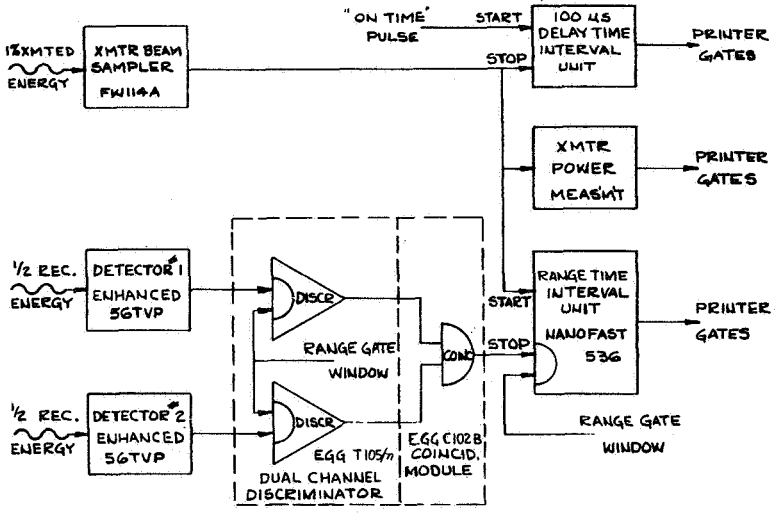


Figure 1–Apollo 11 on the Moon (July 20, 1969).



OPTICAL SCHEMATIC
LUNAR LASER RANGING SYSTEM

Figure 2



RECEIVER & RANGING SECTION

Figure 3

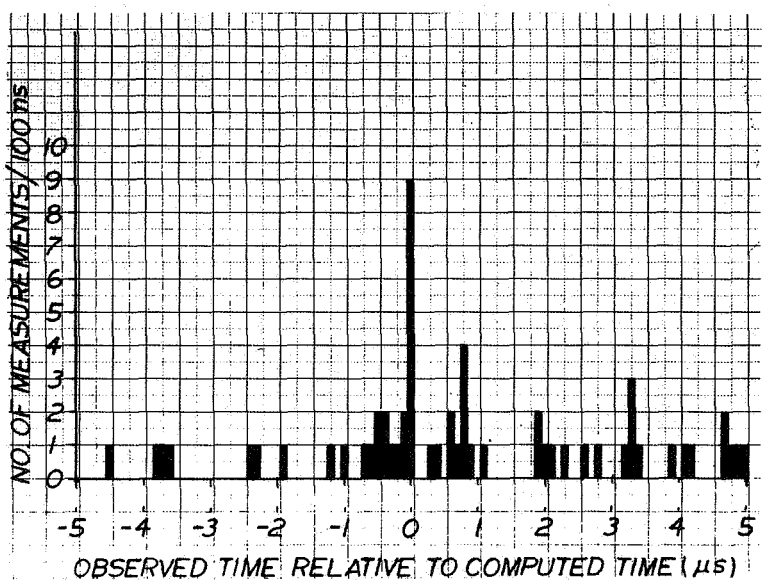

NO. OF MEASUREMENTS/100 ns
0
1
2
3
4
5
6
7
8
9
10
-5
-4
-3
-2
-1
0
1
2
3
4
5
OBSERVED TIME RELATIVE TO COMPUTED TIME (μs)

Figure 4

AUTOMATED TRACKING DATA ANALYSIS

Paul E. Schmid

CHAIRMAN:

Next, a discussion of automated tracking data analysis will be presented by Mr. Paul Schmid.

MR. SCHMID:

The latest developments in computer technology have been applied to the automation of tracking data analysis. Analysis time for the reduction of large data stretches has been reduced from weeks to hours with results displayed on successive frames of microfilm. The purpose of such data analysis is the following:

1. Permit a clear separation of spacecraft tracking-system effects from orbit-computation mathematical-modeling uncertainties.

2. Provide a means for dynamic tracking-system checkout during an actual spacecraft track. (This is in contrast to the usual static collimation tower tests coupled with tracking simulations.)

3. Investigate the character of Earth and other planetary atmospheres by means of tracking system measurements of range, range rate, and angles between Earth-based stations and orbiting spacecraft.

Figure 1 indicates, in simplified form, the tracking data flow that has been completely automated. The GSFC computers utilized are the IBM 360, model 95 or 75, and the Stromberg Carlson 4020. The data from the site are raw measurements of range rate, range, and angles. Such information consists of either Doppler cycle counts or time intervals for a fixed Doppler cycle count with a fixed station inserted frequency offset; time intervals relating total two-way radiowave propagation time, for a measure of slant range; and angle data referenced to a particular local tracking station coordinate system. It is necessary to convert such raw data to metric form in order to relate it directly to the physics of the tracking situation.

After the tracking data have been reduced to the common denominator of range (km), range rate (km/sec), and angles (degrees), they are either converted to a plotting routine that permits a direct film view of metric measurement versus time, or they are subjected to a polynomial curve fit up to 7th order, autocorrelated, and then plotted on film. The direct plot

permits extraction of gross variations in tracking measurement; the autocorrelation analysis provides a means for extracting periodic variations often masked in direct data viewing. It should be noted that this analysis is performed prior to orbit and trajectory computation. This obviates consideration of uncertainties induced in the orbit computation process by imperfect mathematical models of gravity fields, solar pressure, atmospheric drag, and so forth. Thus, all results from the automated tracking data analysis are directly linked to the tracking system and the intervening electromagnetic propagation media (References 1 and 2).

Figure 2 presents two examples of significant results from the automated tracking system analysis. The Apollo unified S-band ground-transmitter acquisition frequency sweep required for phase locking the Apollo transponder is clearly indicated in the automated direct plot of range rate versus time, recorded during acquisition at the Guam tracking site during the Apollo 4 flight (AS-501). This completely automated film frame is based on 190 data points taken at a rate of 10 measurements/sec. This particular sequence provided an inflight check of proper Apollo transponder frequency lock-up performance. The sweep of ±4 km/sec verified that system specifications were met during an actual spacecraft track. This is an example of the dynamic system checkout of the unified S-band acquisition scheme.

The second plot in Figure 2 shows a result of the automated autocorrelation and spectral analysis that is necessary when extracting perturbations that are small in magnitude relative to measurements of spacecraft motion. The spectrum folding is due to the 1-sample/sec rate relative to a nominal 0.4-Hz spin rate. The following results were obtained:

1. The spacecraft spin rate can be accurately extracted from the same range-rate data normally used as input to orbit computation programs (example: 0.415-Hz determined from range-rate measurement compared to 0.414-Hz obtained from Sun-sensor rotation and telemetered to Earth).

2. Detailed investigation of the spectral phenomena (References 3 and 4) led to a mathematical expression for the range-rate bias induced by spacecraft rotation in a phase-modulated tracking system such as the GSFC range and range rate or the Apollo unified S-band tracking systems.

3. The detailed study precipitated by the automated printout also shows (Reference 4) that bias magnitude depends upon whether the spacecraft antenna receives and transmits with the same or with different polarizations. Thus, for the Explorer-type spacecraft, which uses a nominal

148-MHz up-link, the bias is 4 cm/sec as presently configured. If the circularly polarized antennas had been connected to receive and transmit in the same rather than the opposite circular sense, the bias would have been 1 m/sec. Since system resolution of the 148-MHz GSFC range and range-rate system at a 6-sample/min rate is on the order of 5 mm/sec (Reference 5), the above biases are clearly not negligible.

As a result of this analysis, the spin rate bias effect can now be extracted prior to orbit computation. The equation linking spin rate and range-rate bias magnitude in a phase modulated tracking system is (from Reference 4)

$$|\Delta\dot{r}| = \frac{\left|1 \pm \frac{1}{k}\right|}{2} f_s \lambda_t,$$

where

k = transponder turnaround ratio

λ_t = up-link transmission wavelength

f_s = spacecraft spin rate.

The ± sign depends upon relative spacecraft transponder receive-transmit polarizations. Note that this result is independent of the aspect angle. This is an example of isolating a tracking system variation to achieve improved orbit computation accuracy.

Finally, Figure 3 provides two examples of phenomena linked to the radiowave propagation medium, in this case the Earth's lower atmosphere (i.e., the troposphere). Again, automatic plotting is the key to checking significant stretches of data for anomalous behavior. The multipath phenomena picked up in a stretch of GSFC range and range-rate low-angle data show variations of up to ± 4 mrad, well above the angle system electronic servo amplifier noise, which is less than 1 mrad. In this case, however, the result is more important for verifying models of the characteristics of the troposphere and the Earth's reflectivity rather than for improving angle measurements. That is, the multipath phenomena at the nominal NASA frequencies of 2 GHz are associated primarily with the smaller antenna apertures (15-ft diameter or less) where precision angle data are not available simply because of relatively wide antenna beamwidths. Angle data are often used to initiate an orbit convergence (i.e., to perform the first guess or iteration). However, the final orbit computation iterations are invariably based on measurements of range and range rate. Of the latter two metric measurements, range rate appears, for today's technology, to be the most

accurate, since measurements of fractions of a centimeter per second are readily obtained.

The last automated display, surface index of refraction (Figure 3), is not strictly linked to tracking data but rather to meteorological measurements sent to GSFC during an extensive test involving all Apollo tracking sites. This was performed to measure the extent of tracking measurement uncertainty attributed to changes in surface refractivity N_S, a constant contained in all mathematical models for correction of radiowave refraction effects in the lower atmosphere (Reference 6). The results in this case indicated that the Apollo scalar values of N_S, based upon average monthly estimates at each Apollo unified S-band site, were satisfactory in terms of overall tracking system performance. The surface index of refraction is given by

$$n_S = 1 + 10^{-6} N_S.$$

The foregoing are but a few examples of the type of tracking data analysis suitable for automation. The automation goes a long way toward reducing the man-hours required to find out "what is going on." The "why," however, still requires human intervention.

REFERENCES

1. Schmid, P. E., "The Conversion of Fundamental Tracking Data to Metric Form," NASA X-551-69-3, January 1969.

2. Grenchik, T. J., and Putney, B. H., "A Review of Goddard Range and Range Rate System Measurements and Data Processing Techniques," NASA X-551-69-137, April 1969.

3. Marini, J. W., and Murray, C. W., Jr., "Effect of Satellite Spin on Explorer 33 and 35 Doppler Tracking Data," NASA X-551-69-52, February 1969.

4. Marini, J. W., "The Effect of Satellite Spin on Two-Way Doppler Range Rate Measurements," NASA X-551-69-104, March 1969.

5. General Dynamics, "Design Evaluation Report R-67-042, Goddard Range and Range Rate System," December 1967.

6. Schmid, P. E., "Atmospheric Tracking Errors at S- and C-Band Frequencies," NASA Technical Note D-3470, August 1966.

AUTOMATED TRACKING DATA ANALYSIS

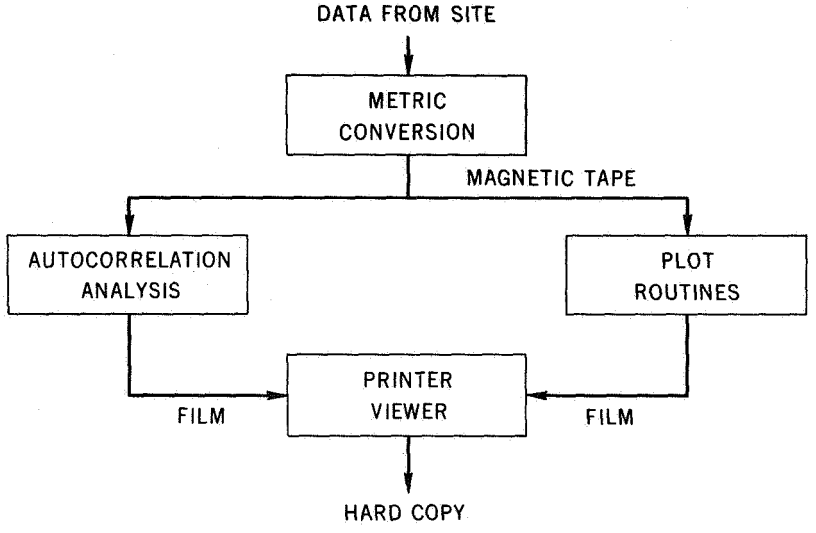
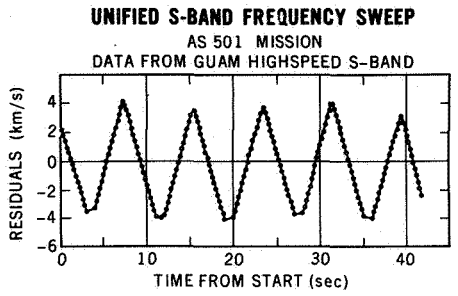


Figure 1

TRACKING SYSTEM ANALYSIS



START OF DATA	SAMPLE RATE	DEG	MEAN	SIGMA	USED	REJ
313 20 12 5.70	.220	5	-2.01×10^{-14}	2.27	190	0

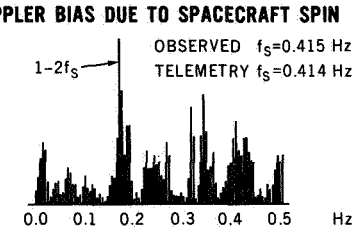
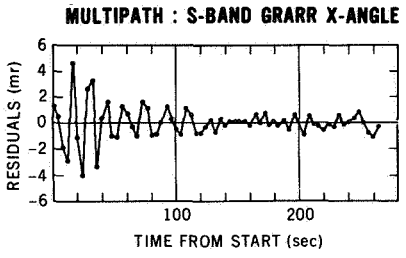


Figure 2

RADIOWAVE PROPAGATION EFFECTS



STATION	START OF DATA	SAMPLE RATE	DEG OF FIT	MEAN	SIGMA
CRO 178 1 4 8.04	4.000	5	3.84×10^{-12}	1.227	

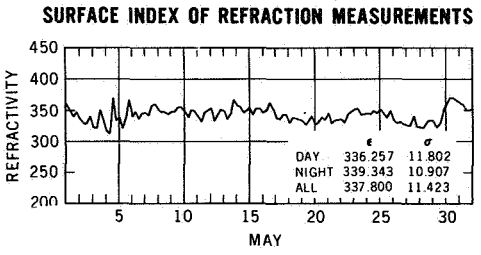


Figure 3

WIDEBAND LOW-NOISE MASER AMPLIFIERS

Wayne E. Hughes

CHAIRMAN:

Our next paper will be given by Mr. Wayne Hughes. His subject is wideband low-noise maser amplifiers.

MR. HUGHES:

Communications systems like those of the proposed tracking and data relay satellites will require ground station receivers with bandwidths of 500 MHz or more so that full-time, full-band coverage can be maintained. To keep spacecraft power requirements to a minimum, the ground station receiver should have the lowest possible noise figure. Since masers are the least noisy of all microwave preamplifiers, we have begun the development, in-house, of a 500-MHz-bandwidth maser preamplifier.

In the design of masers, the bandwidth is determined by the linewidth of the material being used. Good single-crystal materials have linewidths on the order of 60 MHz, as shown on the left side of Figure 1, where the vertical axes are spin energies divided by Planck's constant to give units of frequency. If we require bandwidths greater than about 60 MHz, we must distort or broaden the line in some manner. This can be accomplished by applying a distorted or inhomogenous magnetic field to the material, with the inhomogeneity being either in direction or in magnitude. For example, the ATS maser developed at GSFC used magnetic field magnitude distortion to achieve a 130-MHz bandwidth. However, to extend this technique to a 500-MHz bandwidth is extremely difficult, so we are employing the technique of magnetic field direction distortion.

The distortion could be accomplished in several ways; for example, we could use a long, single crystal with several magnets placed along it with each magnet oriented in a different direction. Conversely, we could use a single magnet with several differently oriented crystals. We use the second technique, but rather than several differently oriented crystals, we use a powder or nonoriented ceramic material. This has the effect of presenting an almost infinite number of differently oriented crystals and produces broad bands in the material, as shown on the right side of Figure 1.

Thus far we have been able to show experimentally that these broad bands do exist in iron-doped rutile, and, by using a combination of experimental and theoretical data, we can construct a theoretical absorption

spectrum for the material, as shown in Figure 2. We are able to identify the major points in this spectrum and to indicate the expected bandwidth and those regions in which we expect maser operation.

To demonstrate the feasibility of this type of maser, we have performed experiments using a multimode cavity in which several frequencies resonate simultaneously. In this configuration, we have obtained maser operation at frequencies separated by 165 MHz at X-band, which indicates that the linewidth is greater than 165 MHz. Thus, using this technique, maser preamplifier bandwidths of 500 MHz are not unreasonable.

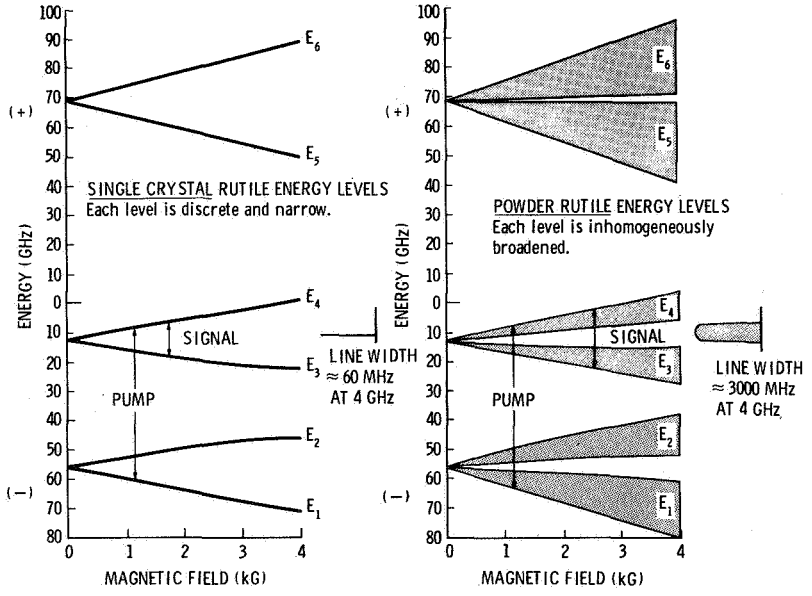


Figure 1–Spin energy levels of single-crystal and powdered rutile.

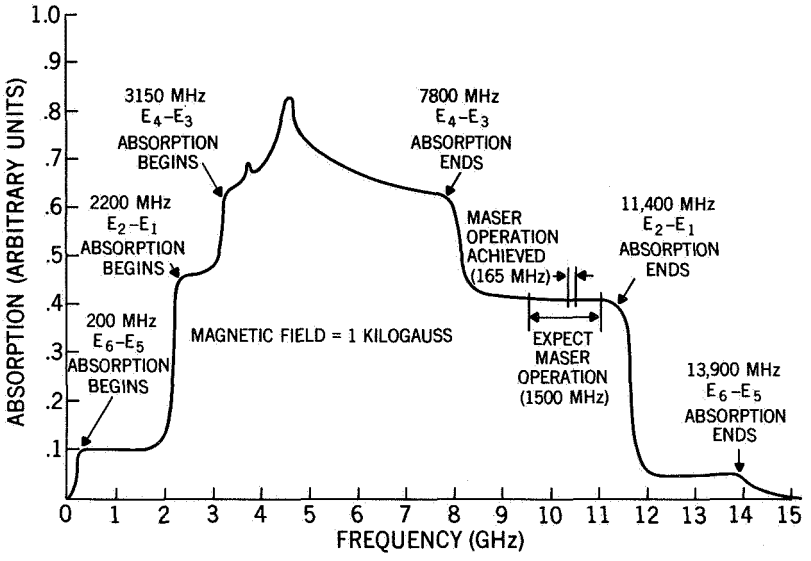


Figure 2–Theoretical absorption of iron-doped rutile.

GSFC FIELD-OPERABLE ATOMIC HYDROGEN STANDARD R&D

Harry E. Peters

CHAIRMAN:

Next, Mr. Harry Peters will discuss GSFC field-operable atomic hydrogen standards.

MR. PETERS:

Time and frequency standards based upon the hydrogen atom, as employed in the atomic hydrogen maser, are the most stable and reproducible devices presently known. Hydrogen maser work at GSFC has as its goal the development of field-operable hydrogen standards that exhibit the exceptional stability and reproducibility characteristics of the hydrogen maser and, at the same time, incorporate all the functional and operational characteristics that make such a standard useful in practical applications.

The first efforts in this field at GSFC, which were begun in 1966, were directed at R&D work to optimize the maser design and to develop the subsidary electronics systems required in applications. Then, in 1967-68, the design and construction of prototype operational standards were begun. During the past year (1968-69), four of these standards have been completed and laboratory performance tests have been made. At present, field tests of these new standards are in progress. Figure 1 shows one of these new units. The standard is mounted on a removable carriage used in transporting it to field stations.

Although hydrogen masers used as laboratory instruments have been made commercially by one company and experimental hydrogen masers are now in use in several university and government laboratories, no practical device has been available that could be routinely used in the field. The present NASA prototype hydrogen standards are designed to fill this need. These standards are an entirely in-house development and are quite different in concept from earlier devices, incorporating several important innovations and improvements.

In the maser itself, an exceptionally large bulb design has been achieved that results in an extremely high oscillating line Q. Improvements in temperature control and cavity isolation have resulted in a very simple, stable, and controllable cavity configuration, and improvements in beam optics and source design have resulted in greatly improved hydrogen

economy. Among the new electronic subsystems in these standards are a unique automatic cavity tuner and a versatile, low-noise receiver-synthesizer system.

These new standards are complete, self-contained frequency and time standards. They include the hydrogen maser and all other auxiliary equipment: automatic cavity tuner, receiver and synthesizer, clock, power supplies, standard output frequencies and time signals, and status monitoring and alarm circuits (Figure 2). Facility for battery operation without functional interruption is included.

Tests to date have shown that these new standards meet or exceed their design goals in terms of stability, accuracy, and reproducibility. Most important for NASA applications and in many other important applications are their stability. Figure 3 shows the results of stability comparisons made among the new standards and compares their performance with that of the best commercial rubidium and cesium standards. Also shown in Figure 3 is a curve giving the present NASA tracking stability requirements, as well as a curve representing the stability required by radio astronomers for very long baseline (VLB) stellar interferometry. Of particular importance for NASA is the fact that the availability of hydrogen standards, such as these new prototypes, will provide a single standard with stability that satisfies both timing precision and frequency stability requirements. The cesium standards currently used have sufficient long-term stability for present timekeeping purposes but are much too unstable for three-way doppler tracking, and of course, the short term instabilities of cesium mean that very long measuring times are required for the setting of clock rates. Although the short term stability of rubidium standards is just adequate for the local oscillator requirements of tracking stations, they cannot be used for timekeeping with the desired precision without frequent resetting of their frequency. The hydrogen standards will exceed NASA's present requirements over the entire range of measuring times and also provide a large margin for future demands for improved stability.

To provide field tests of the hydrogen standards, a number of interesting applications are being pursued. Two units, one at the MIT Haystack Observatory (Massachusetts) and one at California Institute of Technology's Owens Valley Observatory (California), have recently been used as stable local oscillators for a series of very long baseline stellar interferometry experiments designed to test the validity of general relativity's prediction concerning the gravitational deflection of radio waves. In another series of tests, one of the new standards has been taken to the National Bureau of Standards (NBS) laboratories in Boulder, Colorado, and there in conjunction

with NBS scientists, timing comparisons are being made with the NBS ensemble of commercial cesium standards that are used to maintain the NBS time standard. Frequency comparisons will also be made with the precision, long-beam cesium standards that are presently used to define frequency and time intervals.

The final field evaluation tests of the new standards are planned for March 1970. Three of the units will then be deployed at tracking stations prior to the launch of Apollo 13. This will permit a detailed evaluation of tracking system performance using both the present frequency standards and the new hydrogen standards.

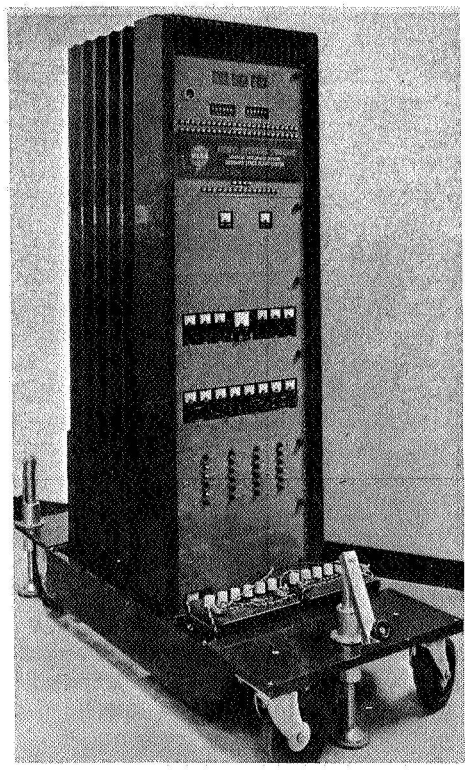


Figure 1

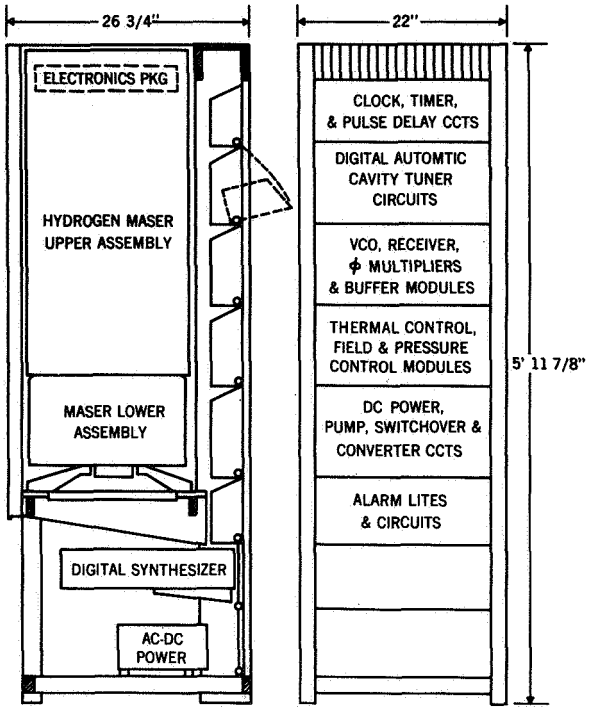


Figure 2

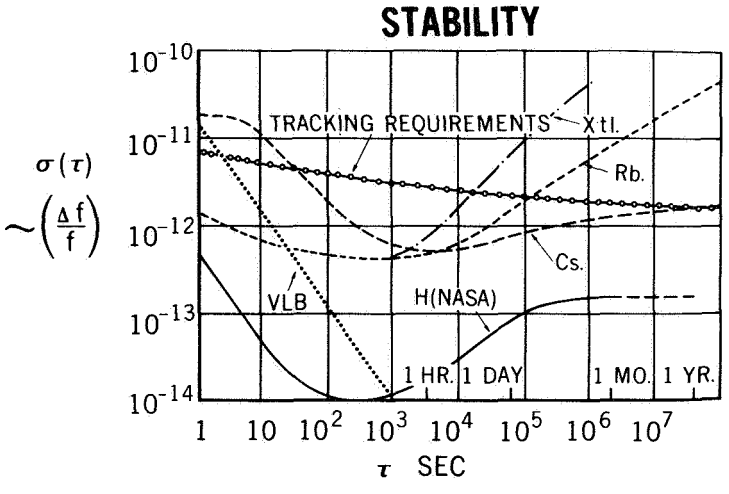


Figure 3

APOLLO SLOW SCAN TV SIGNAL ENHANCEMENT

Edward J. Prokop

CHAIRMAN:

The next contribution is by Mr. Edward Prokop. His topic is real-time TV signal enhancement in the Apollo program.

MR. PROKOP:

Real-time image enhancement of the TV signals from Apollo spacecraft is greatly needed. To provide considerable improvement in imagery quality, image enhancement could be implemented at many different points along the TV signal path. However, the earlier along the path it is applied, the greater is the possible improvement.

In essence, image enhancement expands the range of signal variation of the imagery detail relative to the total excursion of the signal, thus compressing the frequency range of the video signal. This property allows the following–

1. There is much less degradation of the image if enhanced video rather than original video is transmitted along a noisy channel.

2. The enhanced video can be presented on a medium of low dynamic range, such as a cathode ray tube, without significant loss of imagery quality even when the original TV video has a very wide dynamic range.

3. The enhanced TV image is much clearer to the eye than the original image would be.

In the Apollo program, the need for image enhancement is particularly great because of the very strong lighting contrasts experienced on the moon.

To help appreciate the purposes and potentialities of image enhancement, we should recognize that extensive image enhancement is performed within the retina of the eye. In normal viewing of bright outdoor scenes, the eye is able to accept within the same field of view a brightness range of 1000:1 and yet can transmit the image with very high fidelity to the brain along the optic nerve, which probably has a dynamic range of about 20:1. This is achieved by means of complicated retinal interaction effects which enhance the visual image.

Our visual experiences show us that the image enhancement process in the retina is very effective, suggesting that significant improvements in TV and photography might be achieved if similar principles could be applied. Research in this area resulted in a mathematical expression that appears to describe the image enhancement processing in the human retina. The major characteristics of this expression are that it operates in a feedback manner and that it has a mathematical form analogous to the Lorentz transformation of relativity.

This expression was implemented first to provide photographic image enhancement and later to provide TV image enhancement. The resulting process, which is called feedback image enhancement, has been found to provide significantly better results than previous approaches. The results of the photographic implementation of the Sylvania feedback image enhancement are shown in Figures 1 and 2. Figure 1 shows a reproduction of a direct print made from an aerial reconnaissance negative, and Figure 2 shows a reproduction of the enhanced print.

It can be seen from Figures 1 and 2 that the image enhancement process darkens the light areas and lightens the dark areas of the scene, thereby compressing the range of scene brightness. This range compression allows the brightness differences in local areas to be expanded, with the result that the imagery detail can be made to appear clearer. The process also attenuates low-frequency spatial components of the image, thereby emphasizing high-frequency spatial components, with the result that edge gradients in the image are sharpened.

The same mathematical principles employed in the photographic process have been applied to a conventional TV system. A block diagram of the TV feedback image enhancement system is shown in Figure 3. The video signal from a conventional TV camera is fed through a variable gain amplifier, the output of which is controlled and enhanced through a feedback process. The enhancement is achieved by means of the gain-control signal which raises the amplifier gain in the dark areas of the scene and lowers it in the light areas. The gain-control signal is formed by low-pass spatial filtering of the enhanced image. In more detail, the gain-control signal is formed by feeding the enhanced video to a scan converter, consisting of a vidicon TV camera looking at a TV monitor cathode ray tube. The same scan rates are used for the camera and for the monitor, so the scan converter merely acts as a one-scan-field memory. Thus, the enhanced TV image formed in one scan field may be used in the next scan field. The output from the scan converter is fed to the spatial filter, which includes a number of delay lines having a time delay of one TV line. The

signals from these delay lines are summed to provide low-pass filtering in the vertical direction. Filtering in the horizontal direction is provided by a conventional low-pass filter. Thus, the spatial filter provides two-dimensional low-pass filtering of the image. As a point of interest, if the output of the low-pass filter were to be displayed on a monitor, it would resemble an extremely blurred image of the scene.

Thus, it can be seen that the enhancement process operates in a feedback manner; the enhanced image is spatially filtered to form the gain control signal, which in turn is used in forming a subsequent enhanced image. Since this is a feedback process, it would be unstable if damping were not included in the feedback loop. The system damping is inherent, in that the vidicon has a certain amount of image retention, so that the image it provides is a weighted average of the images from several scan fields.

Our analytical and computer studies, as well as the results of the experimental TV and photographic implementations, have shown that feedback image enhancement can provide images of a much higher quality than can other known image enhancement systems. A simple physical explanation for this may be seen from the following considerations. Figure 4 shows simplified block diagrams describing conventional open-loop image enhancement and feedback image enhancement methods. With the conventional image enhancement method shown in diagram (A), the input image is spatially filtered to form the blurred image signal used to control the variable-gain amplifier. With feedback image enhancement, however, shown in part (B) of the figure, it is the enhanced image that is spatially filtered and processed to produce the gain-control signal. Now, the input image from the camera normally has a very wide dynamic range, whereas that of the enhanced image is lower. In the conventional image enhancement method, therefore, an image of wide dynamic range is spatially filtered, but with feedback image enhancement, an image of lower dynamic range is so filtered.

This difference has important consequences in regions of the image close to strong dark-light transitions, the boundaries between shadows and brightly lit areas. At these transition points, strong spatial transients are produced in the blurred image when the input image, with its wide dynamic range, is spatially filtered. These spatial transients obscure imagery detail. However, when the enhanced image is spatially filtered, the spatial transients are much smaller because the enhanced image has a lower dynamic range. Therefore, with feedback image enhancement, imagery detail is presented clearly throughout the picture, even in the vicinity of strong dark-light transitions.

This study was originally directed primarily to the Apollo slow-scan television system because considerable loss of imagery quality occurred in the scan conversion process. However, the use of the slow-scan data rates is being de-emphasized with the progression of the Apollo TV system development toward compatibility with commercial TV data rates. In summary, the primary objectives of this study were to evaluate the TV requirements of the Apollo program and, after a study of possible TV image enhancement techniques, to choose the best approach and make a detailed design study of equipment for implementing this for a manned space-flight network (MSFN) site. A laboratory demonstration of the selected TV image enhancement approach was then to be performed as a feasibility demonstration. Studies of the use of this image enhancement approach, augmented by any compatible video-processing techniques, were also to be made with the aim of providing improved imagery quality during the retransmission of television from a remote MSFN site to a mission control center and in the recording of television data on video tape. Problems inherent in the extension of the image enhancement approach from black and white TV to color were also to be studied.

The significant achievements to date on the program can be summarized as follows:

1. It has been demonstrated that the feedback image enhancement approach is capable of considerably better performance than other known techniques.

2. A real-time laboratory demonstration has been developed which provides high-quality black-and-white television image enhancement at standard TV rates. This approach could be directly extended to form an operational system for Apollo.

3. Studies have been performed to show various ways by which this approach can be extended to provide image enhancement for color TV. However, careful experimentation with the color TV signals employed in Apollo would be needed before the best approach could be determined and its effectiveness evaluated.

4. A preliminary design study has been made of digital equipment that could apply feedback image enhancement to the Apollo slow-scan TV signals.

5. Preliminary design studies have been made of various methods of digital-scan conversion. It has been shown that some of these could

provide considerable improvement over the existing analog-scan conversion equipment being used for the Apollo slow-scan TV signals.

In conclusion, it is evident that our studies and experimental results show that the feedback image enhancement method greatly expands the imagery-detail signal components, relative to the total video amplitude. This produces a video signal that can tolerate a much lower signal-to-noise ratio in a transmission channel without significantly degrading the resultant image.

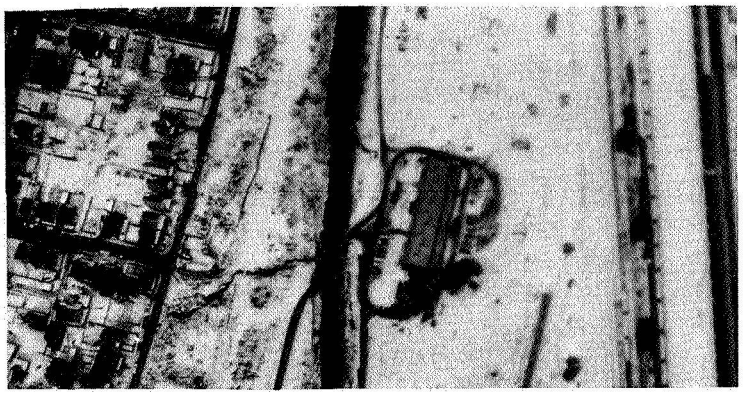


Figure 1

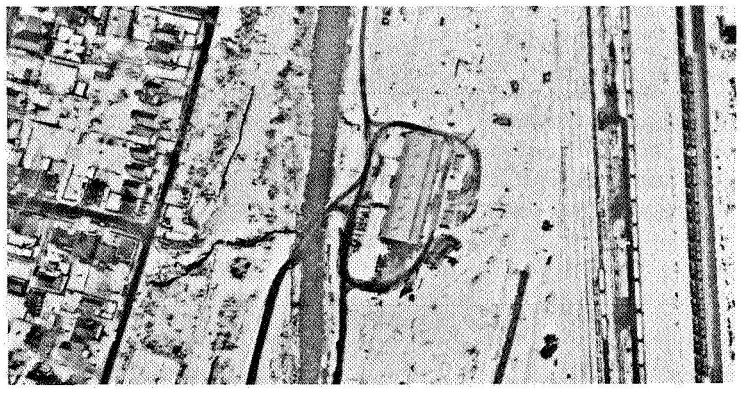


Figure 2

TV FEEDBACK IMAGE ENHANCEMENT SYSTEM

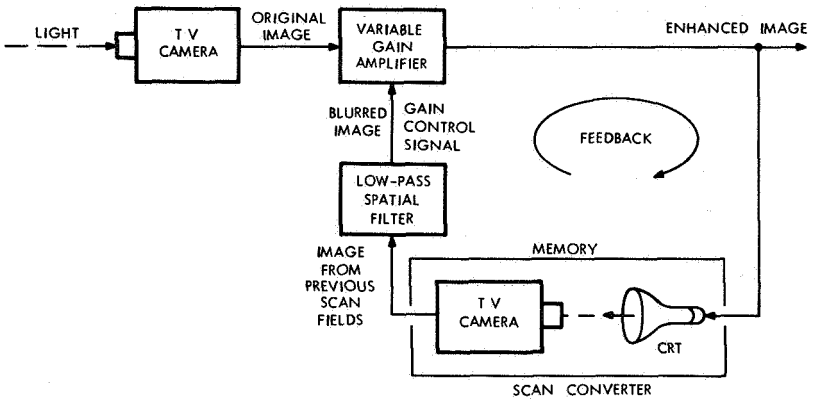
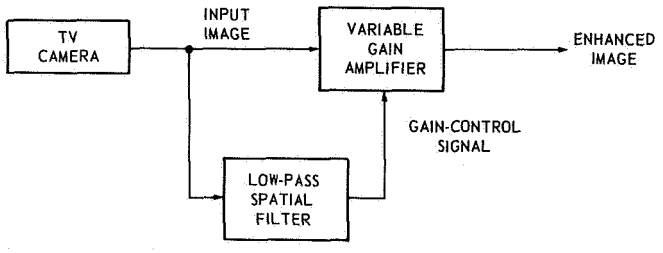
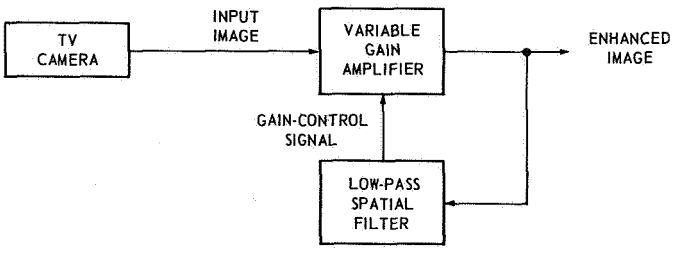


Figure 3



(A) CONVENTIONAL IMAGE ENHANCEMENT



(B) FEEDBACK IMAGE ENHANCEMENT

Figure 4

AN ELECTRON BEAM RECORDER FOR SPACE APPLICATIONS DATA

Harvey Ostrow

CHAIRMAN:

Next, Mr. Harvey Ostrow will discuss an electron beam recorder for space applications data.

MR. OSTROW:

An electron beam recorder (EBR) has been in operation at GSFC for much of the past year. Up to this time, it has been used primarily for recording images from the ATS spin-scan camera since this represents a conveniently available source of high-resolution video information.

The system was briefly described in last year's ART/SRT review. This paper describes the current state of development, indicates the type of performance being achieved, and shows a typical example of an image recorded with the system.

Of the several conceivable ways to increase the overall efficiency of image recording, the most attractive for high-resolution systems is to put the recording film inside the vacuum, completely eliminating phosphors and imaging lenses. It is well known that electrons striking a silver-halide photographic material produce a latent image. Thus, EBR can function as a complete recording system in which the film is directly exposed to a modulated electron stream.

The EBR is shown in Figure 1. Basically, it consists of a high-resolution electron gun; an electro-optical system for focusing, deflecting, and otherwise controlling the electron beam; a film transport mechanism; a vacuum pumping system that maintains the required vacuum in various parts of the recorder; a number of highly regulated power supplies; and other electronic circuits that control the operation of the recorder.

Employing well-established electron microscope techniques, the electron gun, which is at one end of the vacuum system, provides an electron spot approximately 2μ in diameter. The electro-optical system for focusing and deflecting the electron beam uses electromagnetic coils mounted on the outside of the vacuum section.

In its basic concept, electron beam recording can be visualized best by imagining a cathode ray tube (CRT) with the phosphor-coated face

removed and replaced by a photographic film. In this configuration the modulated electron beam strikes the film instead of a phosphor face. It has been found that the reaction of the silver-halide film to an electron beam is much like the film's reaction to light. Thus, the modulated electron beam striking the film creates a latent image that can subsequently be developed by the usual film processing techniques.

A great deal of image recording has been done in the past with a CRT as the source of light. The ease with which such a point light source can be modulated and deflected offers many inherent advantages. However, all CRT recording systems suffer in performance on two counts. First, resolution is limited primarily by the CRT itself and also by the lens and photographic film in the recording camera. Second, wide dynamic range cannot be achieved because of CRT spot halos that are caused by internal reflections of light trapped inside the faceplate and because of phosphor noise that is due to the granular structure of even the most uniformly deposited phosphors. EBR's overcome both of these difficulties and can produce very high resolution recordings on a virtually grainless recording medium with a high dynamic range matching that of conventional photography.

The versatility of direct electron beam recording systems stems from the low inertia of electron beams which can be rapidly and precisely controlled. Thus, EBR systems are particularly useful in image-recording applications requiring not only high resolution and wide dynamic range but also random, variable, or changing scanning rates; controlled geometrical distortion of the recorded images; or rapidly changing spot shapes or spot sizes.

The EBR was developed because of its inherent versatility and its ability to produce high-resolution images of excellent quality. The versatility of the device is an especially important factor. Heretofore, as advanced imaging systems were developed, it was necessary to develop associated recording and display systems. The EBR is already applicable to a number of imaging systems either under development by GSFC or actually in operation, namely, the ATS spin-scan camera, the return beam vidicon camera being developed for ERTS, and a high-resolution image dissector camera proposed for ATS F and ATS G.

The performance characteristics of the EBR are listed in Table 1. Resolution may be defined in many ways. Resolutions of 4000 TV lines at 90% response and 8000 TV lines at 70% response have been measured on a 70-mm format recording. Assuming a Gaussian-shaped response, this represents a limiting resolution in excess of 16,000 TV lines. Scan

linearity of better than ±0.2% in both directions and a dynamic range of 200:1 have been achieved. Densities up to 3.0 can be readily recorded, providing more than 20 shades of gray.

Table 1

Electron Beam Recorder Performance

Characteristic	Performance
Resolution	4000 TVL @ 90% response 8000 TVL @ 70% response
Linearity	±0.2%
Dynamic range	>200:1
Size stability	Better than ±¼%
Frame rates	From 1 frame/sec to 1 frame/min
Line rates	From 3000 line/sec to 1 line/10-sec

The EBR also offers complete flexibility in raster size and format. With minor control adjustments the film format employed can be varied, and the precision with which the size of a particular scan raster can be maintained is nominally one part in 1000. The recorder can operate at line rates varying between 3000 line/sec and 1 line/10-sec and at frame rates varying from 1 frame/sec to 1 frame/min, thereby allowing operation with a potentially wide variety of imaging systems.

Figure 2 shows a typical example of the output of the EBR using video information generated by the ATS spin-scan camera system.

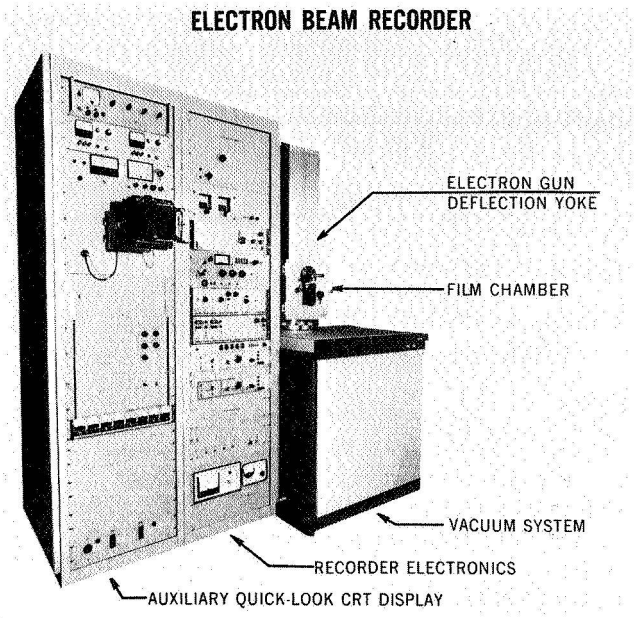


Figure 1

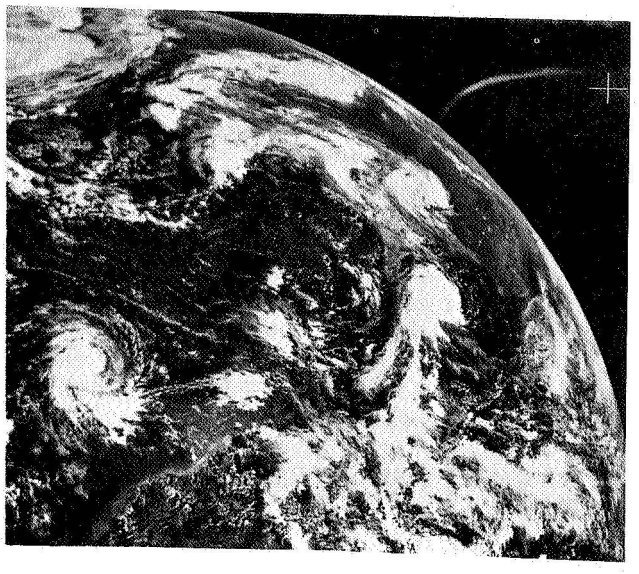


Figure 2

DIGITAL TAPE UNIT TEST FACILITY DEVELOPMENT AND APPLICATIONS

William C. Webb

CHAIRMAN:

We go now to Mr. William Webb, who will discuss the development and application of a digital tape unit test facility.

MR. WEBB:

One of the major missions of GSFC is to process satellite data and to provide the satellite experimenters with their data recorded in a meaningful format on seven-track digital tape. In the process of carrying out this mission, GSFC was confronted with a rather serious tape compatibility problem; due to unrecoverable read errors, experimenter tapes recorded on GSFC tape drives could not be reliably processed on experimenter tape drives. In some instances, tapes recorded on certain GSFC tape drives could not be processed on other GSFC tape drives for the same reason. In 1964, when the tape compatibility problem was first encountered, about 4000 tapes were recorded and distributed to the GSFC satellite experimenters each year. In 1964, the computer time and resources required to reprocess all bad experimenter tapes returned to GSFC and effectively yield 4000 good experimenter tapes per year were available. Based on an anticipated increase to 100,000 experimenter tapes per year by 1970, it was evident we would have to increase GSFC's efficiency of generating good quality experimenter tapes in order to successfully complete our tape mission.

The decision to develop the capability for measuring and controlling the tape recording function of all digital tape drives was justified on the basis of solving the tape compatibility problems described above. At the time the decision was made to develop this capability, there were no satisfactory techniques or equipment available from industry to meet the requirements; therefore, the development was completed as an in-house GSFC project funded under the SRT budget.

The design goals of the digital tape unit test facility (DTUTF) development program were to–

1. Develop the equipment, computer programs, and techniques required to measure and evaluate the performance of digital tape drives by testing their recorded output tapes.

2. Design the routine tape drive testing procedure to be used in measuring and evaluating the internal tape drive performance of the GSFC tape drives. (The test procedure design would make use of the equipment, computer programs, and techniques developed in item 1.)

3. Develop a facility through which GSFC could conduct future research and development work on digital tape drive compatibility problems.

4. Provide GSFC with the necessary techniques and equipment for testing special-problem digital tapes that, because of read errors, could not be processed on computers. (The equipment and techniques would be designed to detect the failure conditions that occurred in the tape drive used to write each of the special problem tapes.)

The DTUTF completed to date consists of a seven-track digital tape drive analyzer equipment and the computer programs necessary to measure the performance of seven-track digital tape drives recording at 200, 556, and 800 bits per inch. The general tape-testing features of the DTUTF are listed below:

1. Character skew.
2. Packing density variations.
3. Parity errors.
4. Recorded signal levels.
5. Inter-record gap (IRG) width.
6. Stray bits recorded within the IRG.
7. Record head asymmetry.
8. Physical tape damage.

These items describe some of the basic recorded data format checks available in the seven-track digital tape drive analyzer equipment. The hardware also has the capability of generating read skew test tapes which contain known amounts of skew inserted into each of the seven tape tracks in a fixed sequence. These tapes can be used to evaluate the amount of read skew and read skew tolerance of digital tape drives connected to computers. The tape drive analyzer can also test new or proposed digital tape drives prior to their procurement by GSFC.

In addition to the tape drive analyzer, a new hardware system has been developed through the fabrication phase which extends the test capability to include nine-track tape drives and adds additional, more comprehensive checks than are currently available in the existing equipment. This system is currently in the checkout phase of development and is scheduled to be completed in March 1970.

The software features currently available in the DTUTF and the function performed by each of the four programs are described below:

1. The read skew analysis program is designed to read the read skew test tapes, generated on the tape drive analyzer, and to compute the amount of read skew and read skew tolerance for each digital tape drive connected to the Univac 1108, CDC-3200, IBM 7010, IBM 7094, or IBM 360 computers.

2. The routine test tape generator program is used to write worst-case data and block length patterns on each digital tape drive connected to the Univac 1108, CDC-3200, and IBM 7010 computers. These tapes are subsequently evaluated on the tape drive analyzer to measure the write performance of each tape drive tested.

3. A set of general support programs was written to evaluate the performance of recorded digital tapes on certain GSFC computers. These programs are used to resolve computer hardware and software problems during the development phase of work on new digital tape data processing programs.

4. Management report preparation programs are currently being developed to interpret the results obtained from the routine tape testing program; to determine the effects of tape drive maintenance and preventive maintenance time allowed, tape brand, and so on, on tape drive performance; and to assist in the preparation of monthly reports on the internal performance of the tape drives.

The DTUTF was first used by GSFC in 1967 to evaluate special tape drive compatibility problems between GSFC and the GSFC satellite experimenters and between tape drives internal to GSFC. Two major tape drive compatibility problems were resolved during 1967 through the use of the DTUTF.

One problem was the inability of the Univac Uniservo VIII-C tape drives to generate good output tapes using Ampex tape. The DTUTF detected a failure on the part of these tape drives to accelerate Ampex tape within the time required by the Univac computer. This tape acceleration problem caused excessive tape data packing density at the beginning of each tape data block. The DTUTF detected the problem and coordinated the effort that eventually solved the problem by modifying the Uniservo VIII-C tape drives.

The other problem resolved was that the STARS Phase-1 telemetry processing line to the CDC-3200 computer tape link was not compatible

because of excessive read skew contained within the CDC-3200 tape drives. The DTUTF conducted tests and recommended corrective procedures which resulted in the solution of this tape drive compatibility problem.

In 1968 the DTUTF began a routine schedule of measuring the internal performance of all tape drives used to support the generation of GSFC satellite experimenter data tapes. Tape drive performance was measured to determine conformance or lack of conformance to industrial tape drive compatibility standards. The standards GSFC uses were established by taking the minimum industrial requirements for tape drive compatibility and adding the required margins necessary to detect and correct tape drive problems, in most cases before GSFC was actually generating bad output tapes. The tape drive compatibility standards used by the DTUTF is accepted by the major computer manufacturers as being a reasonable requirement. All tape drives are routinely tested for conformance to these standards.

Figure 1 contains a plot of the internal tape drive performance and shows the percentage of test tapes that did not meet the GSFC-adopted industry-compatibility standards for each monthly reporting period from February 1, 1968, through October 31, 1969. Three periods were judged to be periods of poor tape drive performance. These periods are marked *X*, *Y*, and *Z*, and the problems detected by the DTUTF and corrective actions required to solve the problems are also described in the figure.

The information in the figure is monitored to determine the effect of tape brand, preventive maintenance time, and so forth on tape drive performance. By inspecting the percentage of unacceptable test tapes detected during test periods 7-12 of 1968 and test period 1 of 1969, one can see the effect of reducing preventive maintenance time. During test period 7 of 1968, the amount of preventive maintenance time allowed was substantially reduced because of operational requirements for computer time. The effect of this reduction was noticeable, as evidenced by the tape drive performance during test periods 8-12 of 1968. The amount of preventive maintenance time allowed was increased in test period 12 of 1968, and tape drive performance returned to a satisfactory level during the subsequent test period.

The figure also shows that once a problem is detected by the DTUTF, as indicated by points *X*, *Y*, and *Z* in the figure, the percentage of unacceptable tapes is reduced to normal very rapidly. The steady downward (improvement) trend in the percentage of unacceptable test tapes demonstrates the benefit of continuing a routine tape-testing program.

Each month, the DTUTF processes approximately 150 special problem tapes submitted for many different reasons from within GSFC. For example, the OSO project has a tape drive compatibility problem between the GSFC SDS 930 and the IBM 360 computer. On the problem tapes written on the SDS 930 tape drives, the DTUTF detected bits recorded in the inter-record gap and short end-of-file gaps. The DTUTF continues to support the OSO project by testing tapes to determine when these two problems are corrected.

Because of tape read errors, the RELAY project could not read digital tapes written in Australia. The DTUTF detected the problem to be due to excessive read skew and provided Australia with information on the corrective action required to solve this problem.

In another case, the IMP 4 experimenter at the South West Center for Advanced Studies was using bad programming techniques which contributed to a high rejection rate for his tapes. The DTUTF detected this and recommended program changes which eventually solved the problem.

GSFC computers are now exchanging output tapes with an effective rejection rate of less than one percent due to all types of tape problems, including software errors, physical tape damage, and tape drive performance. The exact part of the rejection rate due to tape drive performance is unknown at this time, but it is considered to be very good in that something less than the one-percent rate can be from this cause.

The GSFC experimenters were rejecting up to 50 percent of the experimenter tapes written on certain internal tape drives during some periods of 1967. The DTUTF started testing these tape drives for conformance to tape drive compatibility standards in February 1968. The effectiveness of this in controlling the experimenter tape quality is demonstrated by the reduction to an acceptable level of the number of experimenter tapes returned to GSFC. A typical example of the improvement obtained is demonstrated by the tape rejection rate by the OGO 3 and OGO 6 experimenters, located at the University of Chicago, during the period October 1967 through October 1969. During certain months in 1967, the OGO-3 experimenter rejected 50% of the tapes received from GSFC, and on July 21, 1967, GSFC was notified that the OGO 3 data processing had come to a disastrous halt due to 400 unprocessable OGO 3 tapes on hand at the University. The OGO 3 and OGO 6 experimenters at the University of Chicago had an error rate of 10 read errors per 100 tapes for 1968 and zero read errors thus far in 1969. The 10 read errors per 100 tapes for 1968 do not indicate 10 unprocessable tapes per 100 but rather that the quality of the tapes is very good by current standards.

Future plans for the DTUTF include the completion of the new seven/nine-track digital tape drive analyzer by March 1970; the extension, by fiscal year 1972, of the test capability of the hardware to include 1600-bit/inch tape drives; and the development during fiscal year 1971 of the hardware and software required to evaluate the read performance of digital tape drives. Accounting programs to assist the DTUTF in interpreting the large volume of test results will also be written before the end of fiscal year 1972.

Monitoring of the performance of all GSFC tape drives and the processing of special problem tapes submitted from within GSFC will meanwhile be continued as routine services.

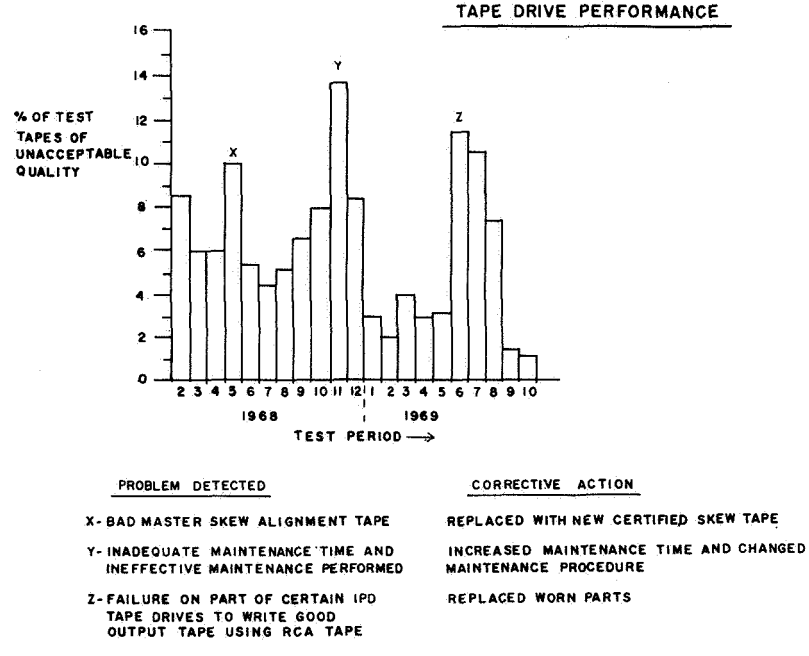


Figure 1

CLOSING REMARKS BY MR. MACKEY

This brings to a conclusion the papers of this session. Before turning it over to Dr. Clark, I would like to thank the many men and women who have helped put this program together. I am not going to try to list them by name, because I am afraid I might miss somebody.

There are, however, three men to whom I would like to give my special thanks: Bill Watson, who arranged and assisted in the displays that you see around the room, and Joe Dunst and Jim Bridger, for yeoman service in planning and organizing the presentation and in helping out with the myriad details involved.

CLOSING REMARKS BY DR. CLARK

Thank you very much, Mr. Chairman. You have done an excellent job keeping us on time all day. There are a few more things that I would like to say.

The first thing is that these reviews never run smoothly just by themselves. I would like to add my thanks to the people who have helped so effectively.

Certainly, first in line are all of the speakers who have done what I would view as the most outstanding job on this sort of thing that I have seen during my 4-year tenure here at GSFC.

Second in line is you, Mr. Mackey, as chairman of today's session, and Dr. George Pieper, the chairman of yesterday's session.

Next, some of the most demanding jobs, on which the excellence of these presentations has hinged more than on any other single factor, were the critiques that were obviously made during the preparation of these talks. Leading this effort have been Mr. Wasielewski, Dr. Pieper, Dr. Meredith, and Mr. Stroud.

This review represents quite an expenditure of time and talent, not only on the part of the men who have done all these things but also by those of you sitting here in the auditorium.

The format is an innovation this year; my enthusiasm is not synthetic. However, GSFC is not the consumer here; by and large, the consumer community is represented here in the audience. So I would like very sincerely to solicit your comments and suggestions. We have here a very distinct innovation, and if this is not the kind of innovation that serves the purposes of the bulk of the users of this commodity, then, no matter how well it may look from the point of view of GSFC, it may not have been a step in the right direction.

If we do not get too many adverse comments, it would be my intent to continue on this road next year. I would hope that we can move, over a period of time, in the direction of essentially wide-open competition at GSFC, from a fabrication innovation or a new management tool, to the more conventional disciplines of engineering and science. I would also hope that we can bring to the attention of the management of GSFC and NASA Headquarters, and of other R&D organizations, both local and distant, all of the

worthwhile work that is done at GSFC in any field. I really mean to have competition across the board.

MR. WASIELEWSKI:

I hope you all realize that in the technological sense this is an experiment, and therefore we need your comments, suggestions, and additions. We think it is worthwhile, but if for some reason you have any doubts, please tell us. Because the format for the technology session is directed towards making your day worthwhile, I hope next year will show a greater attendance and a greater participation.

DR. CLARK:

Thank you, Mr. Wasielewski. I would also like to thank some of the senior people from NASA Headquarters for coming today. I would single out Mr. Vince Johnson from OSSA today, and Dr. Newell and Dr. Naugle yesterday. I was also delighted to see a number of representatives from our sister centers. That is all I have to say, Mr. Chairman.

CHAIRMAN:

Thank you very much, Dr. Clark. And I think the best thing I can do right now is adjourn this session.

NASA-Langley, 1970 —— 34