INVESTIGATIONS OF THE IONOSPHERE BY SPACE TECHNIQUES

S. A. Bowhill

Department of Electrical Engineering

University of Illinois

Urbana, Illinois 61801

Abstract

Much of the impetus to ionosphere research since the International Geophysical Year has come from new types of measurement using space vehicles. The key developments are outlined, together with the contributions that they have made to our understanding of the ionosphere.
1. Introduction

In the mind of the layman, the beginning of the space age was the launch on October 4, 1957 of the USSR satellite Sputnik 1. In one sense this point of view is correct, in that the subsequent 10 years saw an almost explosive growth in studies of the environment of the earth by space techniques. In another sense, however, the nuclei of this growth existed substantially before the launch of the first satellite, and this paper will explore those earlier beginnings and trace their effects on the course of later space investigations.

The usefulness of space vehicles for studies of the atmosphere and ionosphere falls under four main headings:

1. Their ability to transmit or receive radio signals through the ionospheric medium, permitting the properties of the ionosphere to be deduced from the amplitudes, frequencies or polarizations of the signals;

2. The possibility of placing aboard the space vehicle instruments that measure properties of the ionospheric plasma, such as its density, composition and temperature, and the natural electric fields contained within it;

3. The ability to carry to above the atmosphere, and into orbit, sensors and imagers at wavelengths from radio to ultraviolet, that can be used to map the upper atmosphere and ionosphere
and its properties from the space vehicle (remote sensing);

4. The possibility of conducting active experiments on the ionospheric plasma surrounding the space vehicle by applying electric fields, releasing chemicals, or injecting high-energy particles.

After a brief historical introduction, some of the more significant results of space investigations will be described and suggestions will be made as to the directions of future research.

2. Beginnings of ionospheric space research

The threshold of the era of space research can be taken as the five years 1952-1956. In each of these years, an event took place with major consequences for the future of ionospheric investigations; taken together, they form an overture to the intensive period of space research that began in 1957 with the International Geophysical Year.

In 1952 was published the Second Edition of S. K. Mitra's famous text book (MITRA, 1952) which summarized the best understanding of ionospheric characteristics and processes that could be obtained from the optical, radio and geomagnetic observations made from the ground up to that time. One chapter in that book outlines rocket experiments described in unpublished US reports, but these results had as yet no impact on ionospheric theory.

In 1953, a conference on Rocket Exploration of the Upper Atmosphere was held at Oxford, where essentially all the results obtained since 1946 from the V-2, Aerobee and Viking rocket programs in the USA were described (BOYD AND SEATON, 1954). Major emphasis was placed on the development of rocket techniques. This meeting was the ancestor of the series of annual meetings of the Inter-Union Committee on Space Research (COSPAR), the first of which was held in 1960, and whose proceedings entitled "Space Research"
have become the normal mode of publication for recent results of space experiments.

In 1954 the Physical Society Conference on the Physics of the Ionosphere was held at Cambridge. That meeting brought together the best ionospheric research being carried out at that time, primarily from ground-based measurements. Its Proceedings (PHYSICAL SOCIETY, 1955) set a style for descriptions of the ionospheric processes in terms of simple physical models that has continued through to the present day.

In 1955, a conference on Airglow and Aurorae was held in Belfast (ARMSTRONG AND DALGARNO, 1955), with emphasis on the chemical excitation mechanisms rather than on the source of the exciting energy itself. The study of these energy sources by energetic particle sensors carried on space vehicles fathered the new subject of magnetospheric physics, a distinctive product of space research.

In 1956, a conference on Chemical Aeronomy was held at Cambridge, Massachusetts (ZELIKOFF, 1957), devoted to discussions between laboratory chemists and aeronomers on the photochemistry and spectroscopy of the upper atmosphere. Mass-spectrometric measurements of ions and neutral particles from rockets and satellites have given great impetus to this subject in the past ten years.

Though a substantial number of new rocket measurements were made during the International Geophysical Year (1957-8), and the first earth satellites were launched in the same period, space research was never fully integrated into the IGY program. Partly this resulted from extent of advanced preparation needed for rocket payloads, preventing coordinated advanced planning of launch dates and times; and partly from a failure
of ground-based experimenters to realize the far-reaching consequences of space measurements for the interpretation of their data. Following the International Quiet Sun Years (1964-5), greatly improved reliability permitted scientists planning rocket and ground-based experiments to work closely together. The greatly increased scientific yield obtained from the coordinated solar-eclipse campaigns of 1966 and 1970 has now led to the majority of scientific rocket firings being part of one or another combined set of experiments.

In the first scientific satellite programs, experimenters were mainly concerned with developing instrument packages and demonstrating the correctness of the measurements made. Workers were slow to compare their data, or even to reduce data for the same time periods. Newer programs such as the International Satellite for Ionosphere Studies (ISIS) and the Atmosphere Explorer (AE) have recognized this deficiency and have been planned as coordinated programs from the beginning.

3. Radio propagation in the ionosphere

Prior to the availability of space vehicles, there were three regions of major uncertainty in electron concentrations deduced from ground-based ionosphere sounders. These were:

1. The topside of the F2 layer, where it was impossible for radio signals to be reflected;

2. The upper E and lower F regions, between about 110 and 200 km, where it was unknown whether a substantial valley was present in the ionization profile, particularly at night;

3. The D region below about 90 km.

Radio propagation measurements from rockets were particularly suitable for clearing up these difficulties. BERLING (1951, 1954) showed how the error
in a rocket's position, deduced from a two-frequency tracking system known as DOVAP, can be attributed to ionospheric dispersion. From these error measurements he calculated electron concentrations for the E and F layers from several rocket shots. This technique was later elaborated by SEDDON (1953) using a rocket-borne two-frequency beacon transmitter. A related technique used measurements of the delay in a radio pulse transmitted between the rocket and the ground (LIEN, et al., 1954). Two developments have tended to make these types of experiment less useful in modern space research. The first is the availability of reliable in situ sensors for the measurement of electron concentrations (see Section 4); the second is the invention of the incoherent-scatter technique by which the structure of the E and F layers of the ionosphere can be measured in detail and almost continuously.

A second type of rocket measurement involved the Faraday rotation of the plane of polarization of a wave transmitted in either direction on the path between the rocket and the ground. Originally noticed in lunar radar returns (EVANS, 1956), this effect appeared dramatically in some of the first high-altitude missile flights (NISBET AND BOWHILL, 1960). It led to several applications of Faraday rotation in space research: to geostationary and orbiting satellites, discussed below; and to successful measurement of D-region electron-concentration profiles (AIKIN, et al., 1964).

The latter technique involved measurement by an ascending rocket of the directions of the plane of polarization of two frequencies chosen low enough to produce substantial Faraday rotation. This experiment, which amounts to a measurement of the relative phase paths of the ordinary and extraordinary modes, can be successfully combined with measurement of the
differential absorption of the two modes (SEDDON, 1958). As later refined by the use of a null measurement with a spinning elliptically polarized wave radiated from the ground (MECHTLY, et al., 1967), the differential absorption/Faraday rotation technique has given a wealth of information on D-region electron concentrations under a wide variety of conditions, and these are no longer uncertain by more than a few percent, though great day-to-day variability occurs.

A quite different technique was used by CARTWRIGHT (1964) in measuring the phase of a VLF radio signal transmitted through the ionosphere in the whistler mode. He was able to measure electron concentrations below 300 cm\(^{-3}\) at an altitude of 180 km, showing the night-time intermediate E layer.

The rapid horizontal motion of an artificial earth satellite causes the length of its line-of-sight path through the ionosphere to vary rapidly with time. For a radio wave propagated to a fixed ground station the Doppler shift depends on the total ionization along the path; this may be measured if the Doppler shift variation with time is compared with the known satellite velocity (WEEKES, 1958), or if that of two waves of different frequencies is compared (the differential-Doppler technique, ROSS, 1960). Similarly, the plane of polarization of a radio signal transmitted from a satellite appears to rotate rapidly when viewed from the ground, due to varying Faraday rotation in the ionospheric layers (BLACKBAND, et al., 1959).

Both Faraday rotation and differential-Doppler experiments measure the vertical integral of the electron concentration (the total electron content) rather than the ionization profile itself. Similarly, measurements of Faraday rotation of a geostationary satellite signal (GARRIOTT AND LITTLE, 1960) gives the total electron content, but on a more continuous basis than is possible with orbiting satellites.
With the advent of incoherent-scatter sounders, interest in this type of measurement for electron concentrations has lessened, but there is renewed interest in measurements of the scintillation of radio signals from orbiting and geostationary satellites (PARTHASARATHY AND REID, 1959). These scintillations are produced by irregularities of ionization (usually field-aligned) in the E and F regions, associated with various kinds of instabilities in the ionosphere. Measurements of amplitude or phase of a satellite signal on the ground have been used (e.g. by MOORCROFT AND ARIMA, 1972) to determine the shape, size, orientation, and location of the ionization irregularities producing the scintillation, whether they be due to spread F, sporadic E or artificial heating of the ionosphere. Studies of these irregularities are important in communication, as equatorial scintillations have been seen at frequencies as high as 2.3 GHz in signals from geostationary satellites.

A final area in which space research measurements of radio waves have proved exceptionally useful is in the study of whistlers and of VLF and ELF noise. HARTZ (1969) has described the noise environment at satellite altitudes, and GURNETT, et al., (1965) discovered the proton whistler, a phenomenon which is undetectable except by means of space measurements.

In summary, the usefulness of radio-propagation techniques in space research is in the high precision with which the dispersion characteristics in a plasma are known, and the relatively well-understood nature of the diffraction of radio waves by plasma irregularities. In combination with ground-based measurements, these space experiments will continue to furnish useful information.

4. In-situ measurements of ionospheric properties

Probably the greatest contribution of space research to ionospheric physics has been the ability to place various kinds of plasma probes into
the ionized region itself. Those familiar with laboratory plasma measurements know the difficulties associated with outgassing of the walls of the vacuum system; doubts about outgassing from space vehicles led to some initial skepticism about some of the surprising results from in-situ measurements. Most space vehicles, however, move rapidly compared with the thermal velocities of gaseous contaminants (about 3 times as fast for rockets, nearly 20 times for satellites), so instruments on the forward side of the space vehicle are usually completely unaffected by gases carried with the vehicle. Mass spectrometers, the most sensitive of such instruments to contamination, are usually outgassed and sealed at ground level, being unsealed only when the space vehicle is high in the atmosphere.

The first and simplest in-situ sensor was the Langmuir probe (DOW AND REIFMAN, 1949). Because vehicle potential effects were not understood at the time of the first experiments, the first reliable measurements were made with a bipolar dumbbell probe ejected from a rocket (SPENCER, et al., 1962), though later rocket and satellite experimenters have successfully used a variety of probes with much smaller areas enabling the space vehicle to be used as an essentially stable voltage reference (BOYD, 1968). Two major problems have plagued the simple Langmuir probe. Firstly, the theory of electron collection in the presence of a magnetic field is not well enough understood to permit absolute calibration of electron current measurements in terms of electron concentrations. Secondly, electron temperatures derived from Langmuir probes have been found to be consistently higher than electron temperatures measured by incoherent scatter radar. This latter effect seems to be a result of surface contamination of the probe, since probes that are cleaned while in flight show electron temperatures in better agreement with incoherent scatter.
Langmuir probes have been found to be useful for measurement of electron concentrations when supplemented by radio propagation measurements (MECHTLY, et al., 1967). However, other types of probe using the variation in capacitance of an antenna with frequency (JACKSON AND KANE, 1959) or the behavior of the probe impedance in the neighborhood of the plasma frequency (HAYCOCK AND BAKER, 1962) have been found to give accurate results in the E and F regions.

Some of the most important space research results for ionospheric theory have come from mass spectrometric measurements of ion composition. The first reliable ion composition measurements, made during the IGY (JOHNSON, et al., 1958), showed the transition between molecular and atomic ions around 200 km that was predicted by photochemical theory. They also showed nitric oxide ions to be the most abundant species throughout the E region, although the dominant ions produced are those of molecular nitrogen and molecular oxygen. This is now understood to be the result of ion-atom interchange and charge-exchange reactions resulting in nitric oxide as a terminal ion.

Metallic ions were later detected in the nighttime ionosphere (ISTOMIN, 1963) and identified as meteoritic in origin, and were found (YOUNG, et al., 1967) to coincide with sporadic E clouds measured with a ground-based ionosonde, in confirmation of the wind-shear theory of AXFORD AND CUNNOLD (1966). Metallic ions have recently been found at altitudes above the peak of the $F_2$ layer over the equator (HANSON AND SANATANI, 1970), apparently as a consequence of the "fountain effect" suggested by MARTYN (1953).

The upper ionosphere was demonstrated by rocket-borne mass spectrometers (TAYLOR, et al., 1963) to have layers of helium ions and protons floating on the atomic oxygen ions of the topside $F_2$ layer, as required by the theory of multi-ion diffusive equilibrium.
Using continuously pumped mass spectrometers, NARCISI AND BAILEY (1965) found that D region ionization below 80 km consists primarily of hydrated protons. It was later shown (NARCISI, et al., 1972) that D-region negative ions also tend to form hydrated clusters. Though the chemistry of these hydrates is only imperfectly understood at the moment, it appears that they have recombination coefficients much higher than the unhydrated ions, and the degree of hydration may have an important effect on the day-to-day variability of D-region ionization.

The problem of absolute calibration described above also applies to mass spectrometers and to ion collectors measuring total positive ions. In the altitude region where negative ions are not abundant (namely, above about 70 km in daytime conditions) it has been found adequate to normalize the total ion current against the electron concentration; below 70 km, even the neutral flow field around a rapidly moving rocket is uncertain, and published ion concentrations must be regarded with caution.

Measurements of electric fields in the ionosphere are important because of their relationship to magnetospheric convection, auroral precipitation, the dynamo current system and thermospheric circulation. Their direct measurements presents difficulties both from rockets, by reason of their poorly-determined orientation, and from satellites because of the large induced field from the horizontal satellite velocity. Successful measurements however have been made at high latitudes, (FAHLESON, et al., 1971) and more recently (SCHUTZ, et al., 1973) in the $S_q$ current system.

Many of these measurements from space vehicles can now be duplicated using ground-based techniques such as incoherent scatter, though only a handful of installations are capable of measuring electric fields as well.
as electron and ion temperatures and concentrations. There seems little
doubt that combined in situ ground-based measurements will be necessary to
understand in detail the dynamics and chemistry of the ionosphere.

5. Remote Sensing of the Ionosphere

Very soon after the first satellite was placed into orbit, it was
realized that the natural noise level above the ionosphere was probably
sufficiently low to permit the use of a relatively low-powered sweep-
frequency ionosphere sounder in a satellite. Topside sounding of the iono-
sphere was accomplished first from a rocket (KNECHT, et al., 1961) and
later from the Alouette I satellite (WARREN, 1963).

As expected, major advances in the horizontal mapping of the ionosphere
were achieved from this satellite and its successors, including a detailed
picture of the equatorial anomaly and the high-latitude trough. However,
many other important scientific results were obtained (SCHMERLING AND LANGILLE,
1969). These include the storm behavior of the \( F_2 \) layer; ionospheric
irregularities and ducting of high-frequency radio signals; electron concentra-
tions in the low-latitude red auroral arc; the geographical distributions of
light-ion abundance; and the detection of various types of plasma resonances
(see Section 6). It is interesting to note that the number of topside iono-
grams available from the Alouette and ISIS satellites exceeds the total
number available from ground-based ionosondes over their much longer history.

Of great significance to ionospheric physics, though not directly a
measurement of the ionization, is the use of remote sensing of atmospheric
minor constituents by optical techniques. Generally, these include absorption
or emission spectroscopy, or imaging of the horizontal distributions of
auroral emissions.

BYRAM, et al., (1955), in a sunset rocket launch, measured the intensity
of the solar Lyman-\( \alpha \) line and deduced the molecular oxygen concentration in
the D and E regions; occultation of the same line measured from the Solrad-8
satellite (NORTON AND WARNOCK, 1968) has established a seasonal variability in molecular oxygen that is important for explaining the winter anomaly in the ionospheric F$_2$ layer. The same basic technique has been applied with success to the measurement of atomic oxygen, molecular nitrogen and ozone by using various wavelengths.

The 6300-A nightglow emission of metastable atomic oxygen is produced by F-region recombination, and its vertical profile was first mapped from a rocket by HURUHATA, et al., (1966). REED, et al., (1973), using the OGO-IV satellite, have made global maps of the same line that show many features of ionospheric morphology. Nightglow measurements of the 5577-A line of atomic oxygen, produced by the Chapman mechanism, were first made from rockets by BERG, et al., (1956), and satellite studies of this line with a limb scanning spectrometer (DONAHUE, et al., 1973) have shown how E-region atomic oxygen concentrations vary globally.

Many other optical observations from space are important for the physics of the ionosphere, but only two more will be mentioned here. First is the dayglow emission from the gamma bands of nitric oxide. Nitric oxide profiles were first obtained from these bands in the D and E regions by rocket measurements (BARTH, 1964), and the geographic distribution of this photochemically important constituent has been mapped from the OGO IV satellite (RUSCH, 1973). Finally, high-resolution television pictures from the DAPP satellite (PIKE AND WHALEN, 1974) have shown the global configuration of the auroral oval in a striking way.

With the advent of the Spacelab in the next decade, it will be possible to mount much heavier and more elaborate remote-sensing instruments than is possible in small satellites. Many new aspects of ionospheric processes will surely be revealed by these techniques.
6. **Active Experiments from Space**

A field which offers some of the most intriguing possibilities for studies of ionospheric processes is that of active experiments on the ionosphere. For example, the first topside sounder experiment (KNECHT, et al., 1961) showed resonances in the transmitting antenna at a number of different frequencies throughout the HF radio band. Studies of these resonances have led to a much improved understanding of the nonlinear interaction of radio waves with the ionosphere.

The use of rockets to release neutral barium in the E and F regions just after sunset (HAERENDEL, et al., 1967) causes the formation of luminous clouds of both neutral and ionized barium, and from their motion can be found both the neutral atmosphere velocity and the ionospheric electric field. This technique now has wide application both in the ionosphere and the magnetosphere, and is particularly useful when supplemented by incoherent-scatter measurements of ionospheric motion.

Another interesting ionospheric modification is the generation of an artificial aurora by the injection of energetic electrons into the ionosphere from a rocket (HESS, et al., 1971). Later experiments have shown that wave-particle interactions can be studied during electron injection experiments (CARTWRIGHT and KELLOGG, 1974) and that the electrons can be seen to bounce several times between geomagnetically conjugate mirror points (McENTIRE, et al., 1974).

7. **Conclusion**

This brief summary of ionospheric space research has necessarily omitted whole areas of great importance, such as energetic particle distributions, auroral studies and the ionospheres of other planets. It is clear that radio waves have been somewhat eclipsed by direct measurements and optical remote sensing as a space measurement, except for specialized purposes (such as D-region investigations). However, the complexity of the ionospheric processes
now being revealed will require intensive application of all available diagnostic techniques, including ground-based and space radio measurements.
<table>
<thead>
<tr>
<th>Name(s)</th>
<th>Year</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>BERNING, W. W.</td>
<td>1951</td>
<td>J. Met. 8, 175.</td>
</tr>
<tr>
<td>CARTWRIGHT D. G.</td>
<td>1964</td>
<td>J. Geophys. Res. 69, 4031.</td>
</tr>
<tr>
<td>Authors</td>
<td>Year</td>
<td>Journal and Reference</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>------</td>
<td>------------------------</td>
</tr>
<tr>
<td>GARRIOTT O. K. and LITTLE C. G.</td>
<td>1960</td>
<td>J. geophys. Res. 65, 2025</td>
</tr>
<tr>
<td>HÄRENDEL G., LÜST R. and RIEGER E.</td>
<td>1967</td>
<td>Planet. Space Sci. 15, 1</td>
</tr>
<tr>
<td>HARTZ T. R.</td>
<td>1969</td>
<td>Proc. IRE 57, 1042</td>
</tr>
<tr>
<td>HAYCOCK O. C. and BAKER K. D.</td>
<td>1962</td>
<td>Electronics 34, 88</td>
</tr>
<tr>
<td>JACKSON J. E. and KANE J. A.</td>
<td>1959</td>
<td>J. geophys. Res. 64, 1074</td>
</tr>
<tr>
<td>MARTYN D. F.</td>
<td>1953</td>
<td>Phil. Trans. Roy. Soc. 246, 306</td>
</tr>
</tbody>
</table>


ZELIKOFF N. (Ed) 1957 The Threshold of Space, Pergamon Press.

-17-