X-641-68-198
PREPRINT

THE POLAR UPPER ATMOSPHERE AND AURORA

KAICHI MAEDA
GEORGE ISHIKAWA

MAY 1968

GSFC — GODDARD SPACE FLIGHT CENTER
GREENBELT, MARYLAND

X-641-68-198

THE POLAR UPPER ATMOSPHERE AND AURORA

Kaichi Maeda
Goddard Space Flight Center, NASA
Greenbelt, Maryland

George Ishikawa
Meteorological Research Institute
Suginami, Tokyo, Japan

May 1968

Goddard Space Flight Center
Greenbelt, Maryland

CONTENTS

THE POLAR UPPER ATMOSPHERE AND AURORA

Kaichi Maeda
Goddard Space Flight Center, NASA
Greenbelt, Maryland

George Ishikawa
Meteorological Research Institute
Suginami, Tokyo, Japan

ABSTRACT

This paper reviews the following three topics in the physics of the polar upper atmosphere and their latest developments: (i) Winter anomalies of the polar ionosphere and their interpretations, (ii) Aurorae and some related phenomena, and (iii) Noctilucent clouds. In each topic, there are still unsolved problems and controversies that could be settled by direct observations by means of rockets and artificial satellites. The importance of rocket-soundings in the altitude range between two limits, i.e., the one around 40 km for balloons and the other around 100 km for artificial satellites, is emphasized for future developments in the physics of the earth's upper atmosphere.

THE POLAR UPPER ATMOSPHERE AND AURORA

1. INTRODUCTION

The upper atmosphere in the earth's polar regions is the locale of many interesting phenomena and unsolved problems arising from the following terrestrial facts:

1. The earth possesses a magnetic-dipole field whose axis is very close to its spin axis.*
2. The earth's spin axis is not normal to the plane of its orbit around the sun, but is inclined to it at 66.5 degrees.

Because of fact 1, solar corpuscular radiations** have more effect in the polar regions than in the lower latitudes, causing the polar absorption (PCA) events, auroras, and related phenomena in the polar upper atmosphere.

Because of fact 2, the polar regions have prolonged sunlit and dark periods, resulting in large seasonal variations.

It should be noted, however, that the largest energy input from solar radiation (the main source of dynamical activities of the earth's atmosphere) occurs in tropical regions where the diurnal variation dominates. In this paper, the following subjects are discussed; first, the so-called winter anomalies in the ionosphere and mesosphere; second, characteristics of the aurora with its related phenomena and some aeronomical problems in the polar atmosphere; and finally, the latest developments in the studies of noctilucent clouds.

The "polar region" discussed in this paper is not necessarily confined to the arctic and antarctic zones; in some cases it includes geographic latitudes below 60 degrees. Geomagnetic poles (especially the south pole) are displaced

*The dipole center is displaced approximately 340 km from the earth's center toward the direction 6.5°N, 161°E and moving northwesterly at present.

**These should be called "solar plasma," except the high energetic solar protons which are called "solar cosmic-rays." The plasma holds neutrality as a whole by an equal amount of proton and electrons. (Solar plasma has another name, "solar-wind.")

by more than 10 degrees apart from the earth's axis (geographic poles).* This has great longitudinal effect on geomagnetical events.

2. WINTER ANOMALIES

If the solar ultraviolet-radiation is regarded as the sole source of ionization and heating of the upper atmosphere, there are several phenomena which cannot be explained by this concept. These are called "anomalies." (Ratcliffe, 1960).

2.1 F2 Winter Anomaly

The maximum radio frequency reflected by the earth's ionosphere is the critical frequency of the F2-layer, f_0 F2. This corresponds to the plasma resonance frequency at the F2-peak which is given by (Appendix A)

$$f_m \simeq 9 \cdot 10^3 \sqrt{N_m}, \tag{2.1}$$

where N_m is the number of electrons per cc at the F2-peak.

If one ignores the dynamical motion of air in the F-region, the daytime N_m can be obtained by the equilibrium between the production rate due to air photoionization and the loss rate due to collisions (attachment and recombination). It can be easily shown that (Appendix B)

$$N_m \propto (\cos \chi)^{\frac{1}{2}}, \qquad \text{or } f_0 F2 \propto (\cos \chi)^{\frac{1}{2}}, \tag{2.2}$$

where χ is the solar zenith angle.

---

*In general, two types of poles can be defined:

(1) Dip-pole (magnetic pole), the point on the earth's surface where the magnetic field is vertical; (In 1965, their locations are 75.5°N, 101.0°W and 65.5°S, 140.3°E, respectively.)

(2) Geomagnetic pole, which corresponds to the intersection of the centered dipole axis with the earth's surface (i.e. 75.5°N, 101.0°W for north-pole in 1965). For global-scale problems, the geomagnetic pole is more important than the magnetic pole (the latter is greatly influenced by local magnetism in the earth's crust).

This indicates that the maximum values of f_0F2 (or N_m) and the amplitude of its diurnal variation should be larger in summer than in winter. Observed results, however, disagree with this conclusion (Figures 1 and 2).

Note that the maximum values of f_0F2 and its diurnal variation occur in the auroral zone (around 70°N or S) and decrease again in the polar cap region (Figure 1) and both maxima (noon value and amplitude) increase with sunspot activity (Figure 2). These anomalies show further complications, such as universal-time (UT) dependence and hemispheric asymmetry (higher electron density in the southern hemisphere).

These anomalies may be due to dynamical effects such as the following (for which there are several arguments):

1. Trans-hemispheric meridional circulation (F. S. Johnson, 1960).
2. Horizontal wind shear and geomagnetic control (Hill, 1960).
3. Ionospheric current (Sato and Rourke, 1964).
4. Ionization by solar-charged particles.

Effect 1 brings electrons in the summer hemisphere (photo-electrons) by direct flow to the winter hemisphere, where the contraction of large-scale subsidence further enhances the relative electron density. This explains the higher F-region electron density in the polar winter, but not its large amplitude of variation.

Effect 2 can be seen from Figure 3. This effect is combined also with Effect 3 (Duncan, 1962). Apart from dip-pole, the magnetic lines of force are slightly oblique. If two winds flow in opposite directions, electrons accumulate between the two layers of counter-wind (wind shear). This is because electron motion is restricted along the magnetic line, while dominant neutral particles move with horizontal wind. Since winds and currents depend on the local time and season, the above explanations have been accepted so far but have yet to be proved by direct observations of wind or current. Note in Figure 3 that the position of the f_0F2 peak should fit in these two layers and that, if wind directions are reversed, the peak between those layers is reduced rather than enhanced. According to the top-side sounder (Alouette 1, 2, Explorer 22) in the F-region, the winter anomaly is limited to the region below the F2-peak. Above the peak, the electron density is nearly two orders of magnitude higher in the summer polar ionosphere. This can be explained by higher temperature, and correspondingly larger scale heights, above the F2 peak at the sun-lit (summer) pole.

2.2 Winter Anomaly in the D-Region

In the lower ionosphere, electrons are not only forced to oscillate by the radio-wave electric field, but also collide with neutral air molecules in the dense atmosphere, dissipating radio wave energies into the thermal energies of upper-air molecules. In other words, ionization in the lower ionosphere, below 90 km, causes radio-wave absorption.

Absorptions in the D-region are associated with several other geophysical phenomena, which are discussed in the later sections, and have been observed by many different techniques such as (1) continuous measurements of cosmic noise, (2) forward scattering (oblique long-distance communications), and (3) observations by "ionosonde" – the ground-based vertical sounding of the ionosphere, giving f_{min}, the ratio of multiple-echo intensities, etc.

Based on ionosonde data for 1942-1949 at Slough, England, Appleton and Piggott (1954) reported the following findings:

1. Maximum absorption occurs in winter.
2. The amplitude of the diurnal variation is larger in winter than in summer.
3. Maximum absorptions and amplitudes of diurnal variations both increase with increasing sunspot number.

These can be seen from Figure 4 which plots the average monthly noon values of absorption against inverse Chapman function of solar zenith angle χ. (The inverse Chapman function is well approximated by cos χ except when χ exceeds 80 degrees.)

Two lines in Figure 4 correspond to the two values of the solar activities, i.e., R = 200 and 100, respectively, where R is the average annual sunspot number. Figure 5 also shows these anomalies; monthly median values of absorption (in decibel) for conditions of constant solar zenith angle cos χ = 0.25 (i.e. χ = 75° approx.) at Slough are shown for the period 1942-1949.* Note that the constant solar zenith angle, cos χ = 0.25, corresponds to the noon value in winter (at Slough) but to early morning and late evening in summer.

The winter anomaly absorption was first noticed by Appleton and Piggott (1954). Since then, many statistical studies of this phenomenon have been made

*The values of absorption in decibels are reduced to an equivalent frequency of 1 Mc/s (MHz) by an empirical formula.

by different workers (Dieminger, 1952; Beynon and Davies, 1955; Thomas, 1962), and the following additional features have been revealed:

1. Days of high absorption in winter occur in groups, which show no correlation with geomagnetic activity.
2. Occurrence of high absorption indicates a maximum near 55 degrees geographic latitude (over Europe).
3. The winter anomaly is observed over an area of at least 10^6 km^2 simultaneously.

There are essentially two different interpretations for this anomaly: (1) the precipitation of charged particles; and (2) meteorological effects. The correct interpretation will be found in the future by direct observations (rocket-soundings and satellites). Meanwhile, the outlines of these two ideas seem worth discussing.

2.2.1 Anomaly attributable to meteorological effects

Association of the absorption enhancement in the D-region with increasing temperature in the stratosphere has been reported by many workers (Bossolasco and Elena, 1963, 1966; Gregory, 1961, 1965, 1966; Shapley and Beynon, 1965). As an example, Figure 6 shows the non-deviative absorption index A [db(MHz), Appendix B] and temperature of the 10-mb level around Central Europe for two winter periods: 1958-1959 and 1960-1961 (Bossalasco and Elena 1963).

It should be noticed that thermal expansion of the stratosphere does not compress its upper layer but rather reduces upper layer densities including ions in the D-region, causing outflows from the region where the upward expansion has started. Moreover, it has been known that the warmings and the absorption enhancements are not always simultaneous; often there are 2 or 3 days' lag either way. Until we know the cause of stratospheric warming, we cannot distinguish cause and effect (Kenecht, 1965; Belrose, 1963-1965).

As shown by Nicolet (1965), the warming of the D-region can produce the enhancement of nitric oxide (NO) density. Because of its relatively low ionization potential, NO is regarded as a main ion constituent in the D-region.*

* The threshold ionization potentials of N_2, O_2 and NO are 15.6, 12.2 and 9.2 ev, respectively; corresponding threshold wavelengths for photoionization are 820, 1026, and 1345 Å, respectively. The only species that can be ionized by the solar Lyman alpha (Lyα, 1215.7 Å, 10.2 ev) is therefore NO (Nicolet and Aikin, 1960). Recent rocket measurements by mass spectrometer indicate, however, that below 90-km level NO^+ is a minor ion and major ions (over 70 percent) consist of large-mass species such as 37^+, 45^+up to 55^+, etc. (Narcisi and Bailey, 1965)

From the direct rocket observation that indicates the increase of nitric oxygen concentration below the 100- to 75-km region (Barth, 1964), Sechrist (1967) recently developed a theory to explain the D-region winter anomaly by the well-known warm winter-time mesosphere and by occasional temperature inversion in the winter mesosphere (Jones, et al, 1959).

2.2.2 Effects of particle precipitation

The first paper on the D-region winter anomaly presented by Appleton and Piggott (1954) is based on the comparisons of attenuation of a radio wave at constant solar altitude at different seasons (see Figure 5). As pointed out already, therefore, the attenuations are compared at different local times – noon in winter but near sunset or sunrise in summer. As shown in Figure 7, the precipitation of electrons is maximum around noon-time (L = 4 corresponds to geomagnetic latitude 60 degrees).* (The scattered points in Figure 7 must be examined with an eye on the "median" or trend line.)

From these findings, Maehlum (1967) recently concluded that the total rate of ionization in the D-region, q, is the sum of irregular ionization q_I caused by charged particles and regular ionization, q_R. The latter is due to the solar ultraviolet-radiation and depends on the zenith angle of the sun, χ, and the former depends on the local time, t, i.e.,

$$q = q_R(\chi, h) + q_I(t, h)$$

where h indicates the altitude.

---

*A dipole line of force is described by the simple equation $R = L\cos^2\lambda$, where R is the radial distance from the center of the dipole, and λ is the geomagnetic latitude. The maximum value of R; i.e., radial distance in the equatorial plane, specifies the line of force and is called "L-value." Three adiabatic invariants, constant along a given value of L, designate the motion of a trapped particle in the pure dipole field. The magnetic field in the presence of a magnetic dipole is

$$B = \frac{M}{R^3}(1 + 3\sin^2\lambda)^{1/2},$$

where M is the dipole magnetic moment. The L-B coordinates are equivalent to the R-λ coordinates. Because of the direct connection of L-B coordinates with the adiabatic invariants of particle motions, L-B coordinates are the more commonly used to describe the data of trapped particles in the earth's magnetic field, even though the field is not that of a pure dipole (McIlwain, 1961).

In summer q_R predominates giving approximated $q \propto \cos \chi$. In winter q_I (which has strong local-time dependence and day-to-day variation) predominates.

Appleton and Piggott (1954) concluded that strong absorption phenomena could not have been attributed to particle precipitation because they have no correlation with high geomagnetic activity, K_p. However, according to Maehlum (1967), the large positive correlation between the flux of particle precipitation and K_p does not necessarily conflict with a poor correlation between high absorption and particle precipitation: to increase the absorption only needs a flux of high-energy electrons or hard-spectrum electrons. Since most high-flux precipitations consist of low-energy electrons, the precipitation of high-energy electrons is independent of K_p. An unsolved problem is, therefore, the relative importance of q_R and q_I.

2.3 The Mesospheric Temperature in Winter

In winter, that is, the polar night, the stratosphere is very cold; but, the high-latitude mesosphere (above 70-km altitude) is warmer than in summer. From studies of anomalous sound propagations, meteor trails, and especially the wind system above 60 km, this warm polar-night mesosphere was known before IGY investigations began.

From intensive investigations of general atmospheric circulation, Kellogg and Schilling (1951) suggested that the adiabatic heating of polar-night mesosphere is due to large-scale subsidence in the polar region. This warm mesosphere in the polar night was confirmed by the direct rocket sounding during IGY (Stroud, Nordberg, Bandeen, Bartman, and Titus, 1959). In 1961, Kellogg introduced a new heating process, called "chemical heating" by the recombination of dissociated oxygen molecules. This drastically reduced the required velocity of subsidence which had been more than 1 km per day to compensate for cooling from infrared emission by carbon dioxide and water vapor. As pointed out in the previous section, this warm polar mesosphere in the winter hemisphere supports Barth's interpretation of the high mesospheric radio-wave absorption because of the high concentration of nitric oxide in warm air.

Finally, the radiative heat source in the mesosphere has been shown to be the summer polar mesosphere; the heat source in the troposphere, is always in the equatorial region and the corresponding heat sinks are near the polar tropopause and over the winter pole in the mesosphere (Murgatroyd and Singleton, 1961). It is, therefore, dynamical processes, i.e., convection from the summer mesosphere to the subsiding polar winter mesosphere, that causes the anomalous warm mesosphere at the winter pole and the cold mesopause at the summer pole. This dynamical process can be regarded as the thermally driven annual oscillation

of the earth's atmosphere which has been also discussed theoretically by Sawada and Matsushima (1964). According to recent rocket soundings, the vertical temperature distribution in the warm polar mesosphere indicates wide-range daily fluctuation, while those of the cold summer mesopause is steady for many days (Theon, Nordberg and Smith, 1967). The cold summer mesopause is one of the important factors in the occurrence of noctilucent clouds (which are discussed in the last chapter of this paper).

3. AURORAE AND RELATED PHENOMENA

The aurora borealis, one of the most beautiful and fascinating phenomena in the polar upper atmosphere, has been observed and described by many scholars in northern Europe since the early 17th century. The complexity of this phenomenon in connection with solar activities and the earth's magnetic variations has been discussed in various books such as those by Birkeland (1896, 1908, 1913), Alfven (1950), Harang (1951) and Störmer (1955). The discussions are based on visual observations from the ground, laboratory experiments, and on mathematical calculations. Since World War II, research has been greatly advanced by direct rocket soundings and satellite observations. A detailed summary of pre-IGY auroral research is contained in a book by Chamberlain (1961). Further developments in auroral physics made by the use of world-wide records of all-sky cameras, magnetograms, and rocket and satellite observations are described in an extensive article in the Encyclopedia of Physics (Akasofu, Chapman, and Meinel, 1966). The following sections of this paper will discuss some important features of this very complicated phenomenon. For further details refer to the above sources.

3.1 Temporal and Spatial Variations

At about 60 degrees geomagnetic latitude* (just outside the so-called auroral zone**) in the evening of a geomagnetically disturbed day, a faint white greenish glow is visible over the northern horizon around 9 or 10 p.m. (local time). Gradually this glow spreads toward the south, covering the whole sky, which is now

*To avoid redundant arguments, the description in the following sections is limited only for the northern hemisphere unless stated specifically. From symmetry, similar arguments should be applied to the southern hemisphere.

**There is a maximum number of aurorae along 67 degrees of geomagnetic latitude. The number of aurora appearances decreases rapidly above 70 and gradually below 65 degrees of latitude. These annular zones (north and south) around geomagnetic poles, in which auroral activity is most frequent and strong, are called "auroral zones." As will be discussed later, "auroral oval" is more effective in describing individual (temporal and spatial) auroral activity. "Auroral zone" corresponds to the trace of the midnight portion of auroral oval and can be defined statistically.

seen to consist of arcs stretching east and west. The number of arcs increases, drifting towards the west. If the geomagnetic disturbance is not great, the auroral display fades out into the western sky. On the other hand, if a geomagnetic storm has started (which can be easily detected by checking a rapid-run magnetogram), one of the arcs will brighten suddenly around the midnight sector in the auroral zone. Then the fantastic auroral displays begin, spreading over the whole sky. This most exciting phase of auroral activity is called "auroral break-up." Suppose that we were in the western side of this break-up, i.e., in the evening portion of the auroral zone; then we would see a sudden brightening in the eastern sky, and the arcs swiftly approaching. This fast movement of the arcs toward the west is called "westward surge." On the other hand, in the eastern side of the auroral break-up (or in the morning portion of the auroral zone), all arcs are spread in patchy forms, pulsating and drifting toward the east. This is called "eastward surge." The eastward drift is faster than the westward drift i.e. that of eastward surge is of the order of 10 km per sec; while the westward of the order of 1 km/sec. The westward surge moves along pre-existing arcs; in the eastward surge, the auroral system disintegrates into irregular folds and patches. The brightness of these patches fluctuates rapidly and eventually fades into the early morning twilight. These variations differ case by case; sometimes cycles repeat twice or more in a day; i.e., an explosive expanding phase in which auroral activity spreads simultaneously toward lower and higher latitudes, and a recovery phase in which the northern extended edge retreats slowly and only quiet arcs remain. Figure 8 shows the temporal variation of auroral activity observed from a single station as a typical example (Heppner, 1958). Note that the same displays are seen differently from different locations. For example, arcs far from the observing station are seen as a glow because their lower borders are below the horizon. If the arcs are overhead, they appear fan-shaped, radiating from a point that corresponds to the geomagnetic zenith, and form a "corona." The morphologies indicated in Figure 8 should be read, bearing in mind the effect of location on the aspect of an aurora. By combining several pictures taken by all-sky cameras simultaneously along the Arctic circle (Canada, Alaska, and Siberia stations), it is found that the auroral arc forms a very large loop along a line called "auroral oval." Inside the oval, there are the so-called "polar cap aurorae," in which arcs and bands form fairly straight lines from noon to midnight meridian across the geomagnetic pole. These features can be seen from Figure 9, supplied by Akasofu (1964), which indicates the spatial distributions of auroral arcs and their temporal variations and directions of motion. Note that the widths of these arcs are only of the order of 200m,* and that parts of the arcs contract into small circles that are especially bright; these are called "rays."

*The gyroradius of 20-kev electrons at these altitudes is of the order of 20m. Because of collisions with neutral molecules, however, horizontal spread should be larger than this in the earth's magnetic field.

In Figure 9, concentric circles indicate geomagnetic latitudes. Parts (a) and (b) correspond to a quiet state; part (c) shows the onset of auroral break-up; part (d) shows the peak of auroral activity, parts (e) and (f) indicate decaying phases. Directions of auroral movement are shown by arrows and the position of the sun (noon) is the top of each figure.

3.2 Geomagnetic Storms and Aurorae

The relation between auroral activities and geomagnetic variations has been investigated by many workers (Heppner, 1958, for example). Relations between the so-called "geomagnetic bays"* have been examined extensively by Akasofu (1965), using world-wide data from a magnetogram and an all-sky camera. Conventionally, the geomagnetic bays are explained by the electric-current system and its time variations, which are assumed to be flowing in a maximum conductive layer in the lower ionosphere around 100 km (i.e., lower E-region) level. Except for the daytime equatorial ionosphere, these electric currents are not directly observed, and should be regarded as an equivalent current system, because these currents might not necessarily exist solely in a thin layer as is conventionally assumed. The ionosphere current produces a positive bay (horizontal magnetic field increase) if the current flows eastward, and a negative bay if the current flows westward. Akasofu (1965) used these observational findings to propose a new scheme of the equivalent current system in the polar (or "auroral") substorm (also called "auroral substorm"). This is shown in the Figure 10(b) in comparison with the previous model (a). In part (b) the westward current (auroral electrojet) flows along the auroral oval.

These two models of the equivalent current system stand for a maximum phase of polar substorm. The cycle of auroral substorm can be repeated several times in a day during a period of high geomagnetic activity, which consists of a sudden brightening of arcs and explosive expansion of the active area, within 10 minutes or so, and the slow decaying phase that lasts for one or two hours. Breakup (in which all arcs or bands disintegrate into very bright patches) occurs only around midnight. In the previous model (Chapman and Bartels, 1940; Fukushima, 1953; Vestine, 1940 and Nagata and Fukushima, 1952), shown in Figure (10a), it is assumed that the equivalent current (in either direction) flows along a geomagnetic latitude (in the auroral zone) and the equivalent westward current flowing in the nightside is more intense than the eastward current in the late morning (daytime) side. On the other hand, in Akasofu's new

*In a record of magnetometer (horizontal component), a bay-type variation appears for a short period of the order of hours. This is called the "positive bay" if variation shown a maximum and the "negative bay" if variation shows a minimum.

system the equivalent current flows westward along the auroral oval whose nightside position is close to the auroral zone, but the daytime position is shifted into the polar cap. At point A, both current systems produce a strong positive bay, and at B a strong negative in both systems. The new system, however, produces a weak positive bay at C, which is in the evening sector outside of the auroral zone. The previous system produces a weak negative bay at point C.

The above-mentioned current system analysis can be summarized as follows: (Chapman and Bartels, 1940, p. 301).

The storm-time geomagnetic variation, D, can be expressed by

$$D = Dst + DS + Di \,, \tag{3.1}$$

where Dst is the universal storm-time variation attributed to the (westward) ring current in the earth's magnetosphere,

DS is the "disturbance longitudinal inequality," that is, the function of geomagnetic longitude, ϕ , and latitude, λ , consisting of diurnal and semidiurnal components, and

Di is the irregular part of D, dominant in the auroral zone, and is called "polar magnetic substorm" (Akasofu, 1966),

Akasofu's new current system corresponds, therefore, to

$$D \simeq Di \,, \tag{3.2}$$

whereas the previous conventional current system corresponds to

$$D \simeq Dst + DS \tag{3.3}$$

and

$$Dst = C_0 (\lambda) \tag{3.4}$$

$$DS = C_n (\lambda) \sin (n\phi + \Sigma n) \quad n \geq 1$$

$$\simeq C_1 (\lambda) \sin (n\phi + \Sigma_1) \tag{3.5}$$

Note that the pattern corresponding to DS (Equation 3.5) is fixed with respect to the sun and shifts westward with the rotation of the earth. As pointed out by

Akasofu (1966), however, in the polar region a current system resembling SD appears intermittent with a lifetime of the order of 1 to 3 hours.

3.3 Spectral and Aeronomical Aspects

The aurora brightness described in the previous sections (shown in pictures of arcs taken by an all-sky camera) is mostly due to emissions of the forbidden line (5577 Å), a transition from metastable oxygen atoms, ($^1S_0 - {}^1D_2$) which appears as a greenish white glow. The vertical distribution of this brightness is approximately a Chapman function with a peak around the 100-km level; i.e., its intensity decreases nearly exponentially with altitude above the maximum and decreases very rapidly below it (the layer of maximum brightness). There are red aurorae that have been classified into two types: one, "Type B," appears at a very low altitude mostly at the bottom of bright arcs, as a purplish red border. The emission of type A aurorae consists of two lines from a neutral oxygen atom OI (6300 Å, $^1D_2 - {}^3P_2$ and 6320 Å, $^1D_2 - {}^3P_1$) (see Figure 11) and of the first negative band* from an ionized nitrogen molecule, N_2^+ (3914 Å, $B^2\Sigma_u^+ - X^2\Sigma_g^+$).

Since these red lines from the neutral oxygen atom are due to the transition from a very long-life metastable (110-second) state, $^1D_2$, to the ground states $^3P_2$ and $^3P_1$, heights of emission are limited above 200 km where deactivation** by collision is negligible. On the other hand, Type B emission is mostly due to the first positive groups of neutral nitrogen molecules (5854 Å ~ 10440 Å, $B^3\Pi_g - A^3\Sigma_u$) (Figure 12). The units of brightness, energetic relation, and energy dissipations are discussed in Appendix C. Figure 11 indicates main transitions in oxygen and nitrogen atoms for neutral, [OI], [NI] and singly ionized ions [OII], [NII]. Figure 12 indicates main transitions for molecular oxygen (O_2, O_2^+) and nitrogen (N_2, N_2^+), respectively (Davis, 1965).

Since the height of maximum intensity of Type B aurorae is around 80 to 90 km, energies of the incident primary electrons that are known as the source of this type of aurora should be higher than 200 kev. Although the brightest

---

*The names of positive and negative bands (or groups) originate in cathode-ray or glow-tube experiments. Spectra of the glowing columns near the anode and cathode are called positive and negative groups, respectively. These correspond to the emissions from neutral and singly ionized ions, respectively.

**Transition to other states without photo-emission is called "deactivation," where the energies are balanced by colliding particles (molecules, atoms, or electrons) instead of photons.

aurora is due mainly to the forbidden OI (5577 Å) line, there are many permitted lines from neutral oxygen, neutral nitrogen, and ionized nitrogen atoms. None of these lines is however, stronger than the forbidden line.

Note that the above-mentioned visible emissions are all due to auroral electrons impinging from outer space on the earth's polar upper atmosphere. There are also visible light emissions due to auroral protons; these emissions can be classified in two categories: (1) hydrogen emission and (2) PCA or polar-glow aurorae. The hydrogen emissions observable from the ground are essentially Balmer lines Hα (6562.8 Å), Hβ (4861.3 Å) and Hγ (4340.5 Å), which appear in the quiet homogeneous arcs during the early stages of auroral activity at night. Although the total intensity of hydrogen emission is far weaker than that of the electron-induced aurorae, the discovery of Doppler shift in the Hα profile by Meinel (1951) brought great progress in auroral physics. Doppler shifts of Balmer lines toward the geomagnetic zenith direction are toward the violet side and asymmetric, whereas it is fairly symmetric in the geomagnetic horizontal direction. The asymmetric Doppler shift gives the velocity of impinging protons, which is of the order of 3000 to 4000 km per sec. This is consistent with those energies estimated by the range-energy relation of protons in air, i.e., around 300 kev to penetrate to the depth around 100 km.

The following two reactions can be regarded as the sources of the excited hydrogen atom, H*, which emits Balmer lines:

$$H^+ + H \rightarrow H^* + H^+ , \tag{3.1}$$

$$H^+ + O \rightarrow H^* + O^+ , \tag{3.2}$$

where H^+ is the primary auroral proton entering the atmosphere.

The neutral atomic oxygen, O, is more than a million times abundant than the neutral hydrogen atom, H, around the altitude 100 km, and since the ionization potential of oxygen atoms (13.62 ev) is very close to that of hydrogen (13.60 ev), the charge exchange (resonance-type) of auroral protons with atmospheric oxygen atoms (Equation 3.2) is a main source of excited-state hydrogen, H*, the emitter of Balmer lines in aurorae. Above 500-km altitude, another charge-exchange reaction (Equation 3.1) becomes important, which has been known as a possible source of the ultraviolet Lyman series. The cross sections for reactions 3.1 and 3.2 are very large below 10 kev (both of the order of 10^{-15} cm^2), but decrease rapidly with energy above 10 kev (less than 10^{-18} cm^2 at 200 kev). Emission of hydrogen lines occurs, therefore, near the end of the range of incoming protons; i.e., the emission dominates below a 100-km level (Eather, 1967).

Note that these temporary neutralized auroral protons spread horizontally more than several hundred kilometers across the earth's magnetic field before being ionized again or stopped by colliding air molecules. Davidson (1965) has shown by means of Monte Carlo type calculations that these spatial spreads of auroral protons occur in the polar upper atmosphere. The results explain why auroral hydrogen emission does not have a fine structure such as arcs or bands produced by auroral electrons. Note that the hydrogen emission lines are not only in the visible region, but also in the ultraviolet (Lyman) and in the infrared (Paschen, Brackett, etc.). Because of the atmospheric ozone, the Lyman lines are not observable from the ground; they have been measured by rocket sounding.

Another proton-produced aurora – the "polar glow aurora" – is a vast air glow occurring simultaneously with so-called "polar-cap absorption" (PCA).* According to Sanford (1962), the maximum brightness is comparable to that of the Milky Way, although strong emissions are essentially the same as the type B (purplish red) aurora.** This similarity to Type B aurora is predictable, because the polar-cap glow is due to the direct excitation and ionization of molecular nitrogen and atomic oxygen by the solar protons which penetrate the atmosphere below the 100-km level, (proton energy is 5 to 100 Mev).

Figure 13 shows the example of the polar glow aurora reported by Sanford (1962), indicating the time variations of auroral glow, with cosmic-noise absorption (riometer data at Thule), K-index at Scott station and Kp-index.

3.4 Auroral Infrasonics

The large-scale auroral displays produce sound waves by disturbing the earth's tenuous upper air.*** These waves have frequencies below audible range

---

*After the eruption of flares on the sun's surface, mostly within a few hours, signals of the high-latitude-path radio wave disappear. This is called "polar cap blackout" and has been explained by D-region ionization enhanced by high-energy solar protons (5 to 100 Mev). By means of a "riometer" (a relative ionosphere opacity meter that measures relative-intensity variations of cosmic-noise around 30 MHz), details of D-region ion density variation have been studied and named "Polar Cap Absorption."

**During intense auroral activity, auroral bands appear across the polar cap, as mentioned before. These bands are produced by electrons, whereas the polar cap glow (a weak glow covering whole polar cap area almost uniformly) is produced by protons. Thus, both are clearly distinguishable.

***Audible sounds during auroral activity have been reported by many observers. These seem rather local phenomena, if it is real, probably due to the electro-static discharge near the ground through the dry, cold air. (Chapman and Bartels, 1940, p. 466)

and are called "infrasonic waves" or "infrasonics." Since there are many kinds of pressure waves and disturbances in the atmosphere, especially near the ground, these subaudible sounds with intensity of the order of several dyne/cm^2* should be listened to with the aid of the following device:

In order to eliminate non-periodic local disturbances such as the pressure fluctuation due to wind, a long pipe with many holes spaced at random intervals is attached to each detector's input (microphone). Local pressure fluctuations whose scale is less than the length of the pipe can be cancelled out; only large-scale waves can be detected. In order to find the direction of propagation of the traveling pressure wave, more than three of these detectors are installed approximately 10 kilometers apart. The differences between the times when the wave reaches the detectors indicate the normal direction of the traveling wave front.

Figure 14(a) indicates the diurnal variations of arrival direction of infrasonic waves observed at the National Bureau of Standards (NBS) stations** in Washington during magnetic storms, as measured by the above procedures. As described in section 3.1, auroral activity shifts along the auroral oval within a day, and the most active part, (i.e., the auroral breakup occurring around the midnight sector) moves essentially along the auroral zone clockwise around the geomagnetic pole. Three cases are shown in Figure 14(b), corresponding to three different local times in Washington*** (21, 0, and 3 o'clock, respectively); the hatched area in each case indicates the position of the most active part of the aurora. The location of Washington is also shown by an asterisk in front of an arrow. The time-variation of the arrival direction of the traveling pressure wave (infrasonics) during a magnetic storm observed in Washington agrees with the assumption that these waves originate in the active part of the aurora. Since the pulsating aurora appears just after the auroral breakup (Figure 8) it was thought that the source of the waves might be the pulsating aurorae caused by fluctuating ionospheric heating from the precipitations of auroral electrons occurring at the most active phase of aurorae (Maeda and Watanabe, 1964).

According to the recent observations performed inside the auroral zone, the sources of these infrasonics are not pulsating aurorae but the supersonic horizontal movements of auroral arcs (Wilson and Nichparenko, 1967). As described in the section 3.1, auroral arcs move rapidly westward in the evening sector (i.e., the west side of the breakup) and shift eastward in the morning sector (i.e., the east side of the breakup). This is also shown in Figure 9, where

*Of the order of one millionth of atmospheric pressure on the ground.

**The details of equipments and locations of stations are reported by Chrzanowski et al. (1961)

***75° WMT.

the arcs shaped like a capital omega spread both westward and eastward around the midnight section, (i.e., the breakup portion), constituting the westward and eastward surge. As mentioned in section 3.1, the eastward surge is generally faster than the westward; both are generally supersonic. Thus, shock waves can be formed and propagated in the normal direction of these arc-fronts. Figures 15 and 16 show an example of the detailed correspondence between the appearance of auroral infrasonic waves and the supersonic auroral surge. Figure 15 shows the temporal change of the position of the auroral arc, projected on the map at 1-minute intervals (the time indicated at each arc-position is Alaskan Standard Time (AST) – the local time at 150°W). Figure 16 shows infrasonic records corresponding to this movement of the auroral arc; this is the superposed record obtained at three stations near College, Alaska. The time marks in the horizontal scale are 1-minute intervals; numbers are in UT (UT = AST + 10). A vertical bar at the right indicates the scale of 5 dyne per cm^2 ; ϕ and V are the azimuth of wave normal (measured from north clockwise in degrees) and the horizontal velocity (in m per sec), respectively. Note that ϕ and V are determined by shifting the records with respect to each other until the clear wave fronts in each record fit as shown in Figure 16, where time differences are omitted because they are of the order of several seconds. Since V is the horizontal projection of wavefront velocity (i.e. $V \simeq c \sec \theta$, where θ is the angle between wavefront and the horizontal and c is the velocity of sound), the altitude of the source can be deduced from these observations. The results agree with the assumption that the altitude of the source is around 100 km (the altitude of the bottom auroral arcs). Note also that the horizontal motion of auroral arcs and bands does not mean the motion of a real mass of glowing air, but a visual effect corresponding to the horizontal shift of frontal area in the upper atmosphere when bombarded by the auroral electrons. In this respect, the essential mechanism for pressure-wave generation is auroral heating (Maeda and Watanabe, 1964; Cole, 1965) rather than the ion-motions and their collision with central air caused by the electrodynamic drift of electric current in the lower ionosphere such as the "Pedersen drift" (Piddington, 1964). Since the origin of auroral electrons is not known, the mechanism of supersonic horizontal drift of auroral arcs is not well understood. Another unanswered question is causes of the 20- to 80-second periodicity in the pressure wave, if rapid horizontal movement of arc (i.e., westward and eastward surge) causes the auroral shock wave that decays into an infrasonic wave in the lower atmosphere.

4. NOCTILUCENT CLOUDS

Another interesting and beautiful phenomenon in the polar atmosphere is the occasional appearance of tenuous, silvery blue clouds in the dark rim of prolonged twilight in the summer sky. Since the unusual height of these clouds was

disclosed by Jesse (1896) these mysterious clouds, now called "noctilucent clouds,"* have been studied by many geophysicists (Vestine, 1934; Störmer, 1935; Paton, 1949, 1950, and 1954). Only recently, however, has the mystery of these clouds been unveiled by world-wide general observations and by direct rocket observations (Fogle, 1962, 1964, and 1965; Paton, 1964, 1965, Fogle, Chapman and Echols, 1965; Soberman, 1963; Witt et al., 1964).

4.1 Characteristics

4.1.1 Features indicating atmospheric motions

One of the striking features of noctilucent clouds is their great altitude, which has been measured more than 500 times over the past 80 years at various places in the northern high latitudes, and is known to be nearly constant around 82 km. These clouds generally have interwoven billowy patterns, also wave motions, that have been used to determine the thickness of the clouds although they are very tenuous. According to Witt's measurements (1964) the clouds are about 1 km thick, or less; the vertical amplitude of wave motion is around 2 km and occasionally 5 km. Measurements of cloud drift velocity have been reported since 1885. The most reliable determinations of drift velocity and wave motion have been made by time-lapse photography and motion picture technique (Fogle, 1966). Fogle has summarized old measurement records and his own observations are as follows:

1. Drift velocity is mostly of the order of 50 m per sec with occasional high winds up to 200 m per sec.
2. Drift is generally southwest in the northern hemisphere; northwest in the southern hemisphere.
3. Wavelength (horizontal) ranges from 5 to 100 km with a preponderance maximum around 10 km.
4. Wave propagation velocity derived from measurements of wave-crest movements, i.e., band movements, is more than 200 m per sec relative to the drift of the cloud system.

*According to Fogle (1966), the name "noctilucent cloud" derives from O. Jesse, whose systematic observations of these clouds date from 1885; he named them "Leuchtenden Nachtwolken." This should be distinguished from "nacreous clouds" or mother-of-pearl clouds which appear at an altitude of 15 to 35 km during the winter.

Note that the noctilucent clouds are observed mostly around midnight. The above-mentioned drift motion of clouds should therefore be regarded only as the wind at the polar summer mesopause around midnight.

Haurwitz (1964) interprets short waves as interface waves and long waves as internal-gravity waves. Propagation of the internal gravity wave is not strictly horizontal and its oblique propagation tends toward vertical with decreasing velocity (both phase and group velocity in the horizontal direction approach zero) as the wavelength decreases (Hines, 1960; Eckard, 1960; Maeda and Young, 1966). On the other hand, for waves longer than 10 km to constitute the interface wave* requires a very large vertical wind shear (or large discontinuity in the vertical distribution of temperature). It seems most plausible, therefore, that short waves below 10 km appearing in the noctilucent cloud are the interface waves due to wind shear around mesopause, and the longer waves are the internal-gravity waves. To confirm these interpretations requires further direct measurements of wind and of temperature distribution around mesopause.

4.1.2 Optical features

Since the discovery of their great height, which is far above a balloon's ceiling altitudes, the formation of visible clouds in such a high, tenuous atmosphere has been a mystery; many meteorologists have wondered whether their nuclei are dust or ice.

Spectroscopic observations have shown that the spectra of noctilucent clouds is continuous with no emission lines, with stronger intensity in the blue than in the red and have several Fraunhofer lines. Although these observations indicate that the size of cloud particles is comparable with the wavelength of visible light, more definite estimation of particle size has been made by polarization-measurements of scattered light from the cloud. In Figure 17 the percentage polarization, P, of blue light – 4900 Å observed at several occasions from the ground is plotted against the scattering angle, ψ (the angle between the directions to the scattering point in the cloud and to the sun – Witt, 1960, a and b). Two kinds of theoretical curves are also plotted for comparison; one is based on Rayleigh-scattering which is valid when the size of scattering particle is smaller than the wavelength of scattered light (shown by full lines); the other is based on Mie's theory, which depends not only on particle size but also on particle refractive index (shown by dashed lines). With (1960 b) made a similar calculation, assuming that the particle is not coated by a simple dielectric substance. The results of the latter are quite different

*This is also called the "Helmholtz wave," which can be seen in billowy clouds in the troposphere (mackerel sky).

from observed values (Figure 17), especially at scattering angles exceeding 70 degrees. Comparison of the observed values with Mie's theoretical values therefore leads to the conclusion that the particles of noctilucent clouds are dusts of sizes around 0.1μ, coated with ice, making a total radius of 0.2 to 0.5μ. Finally, the most decisive observation, which settled these arguments was made from direct samplings of cloud-particles obtained by firing rockets into the clouds (Soberman, 1963; Witt et al., 1964). Two complete samples were obtained; one from visible noctilucent clouds, and the other from the same altitude but in the absence of clouds. Electron microscopic analyses* of these samples indicate that nickel and iron particles of the order of 0.1μ diameter are main constituents of the cloud-nuclei; they are coated by ice (possibly hoar-frost). This is confirmed by the calcium film indicator of moisture used with sample collectors on rockets. These results agree, therefore, with the estimations based on polarization measurement and its theoretical analysis. It is also found that particle concentration is at least a thousand times higher in the presence of clouds than in their absence (roughly 10 particles per cc as against less than 0.01 per cc). The size and substances of the cloud nuclei suggest that these are of extra-terrestrial origin.**

4.2 Theory of Cloud Formation

The noctilucent clouds are observable only in the polar summer twilight when the sun is below the horizon with the solar depression angle α (Figure 18) between 6 and 16 degrees*** (Fogle and Haurwitz, 1966; Paton, 1964; Fogle, 1966).

---

*Further measurements of characteristic X-ray spectra are made, exciting the nuclei by an electron beam of 10 to 30 kev; also, neutron activation analyses are applied, confirming that cloud nuclei are of iron-nickel-meteorite origin.

**Details of rocket sampling and various types of analysis of the collected sample are reported in the six successive papers in Tellus Vol. 28, 1964, from page 84 to 117. Further discussion based on the size spectrum of dust nuclei is presented by Webb (1965, 1966 a) who concludes the dust is of meteorite origin. This conclusion agrees with the latest analysis of cosmic dusts and zodiacal light (Singer, 1967).

***As can be seen from Figure 18, if $\alpha < 6°$, the background sky is too bright for the cloud to be seen from the ground P; if $\alpha > 16°$, the sunlight reflected by the cloud is too weak to be seen. The earth's shadow which forms a dark rim of the cloud is not that of the solid earth but of approximately the surface of 30 km in the atmosphere, as shown by H_{sc} (Fogle, 1966).

The shaded areas in Figure 19(a), (b), and (c) – corresponding to geographic latitude, λ, of 40°, 60°, and 70°, respectively – show that the observable period of noctilucent clouds given by the above-mentioned geometric conditions varies markedly with λ and the season. These figures indicate that the noctilucent clouds could be observed over a whole year at latitudes below about 60 degrees. However, Fogle (1966) has shown from his own surveys of IGY-network reports that noctilucent clouds occur most frequently at about $\lambda = 60°$; negligibly below 40° or above 70°.* On the other hand, analyzing the available data of upper air soundings, Murgatroyd (1957) found the lowest temperature (at 80 km in summer high latitudes) to be lower than the radiatively predictable temperature under weak but continuous insolation in the polar summer. As shown in Figure 20, the cold-in-summer and warm-in-winter mesosphere has also been observed by rocket-grenade soundings at Fort Churchill ($\lambda = 59°$) (Stroud et al., 1960). According to the latest observations, the summer polar mesosphere is as cold as 135°K (Nordberg et al., 1965). By calculating global atmospheric circulation (which is consistent with observed wind systems in the upper atmosphere) Murgatroyd and Singleton (1961) have shown that there is a continuous ascending flow in the summer polar mesosphere, at 50 km to 80 km altitude, of the order of 1 cm per sec in the regions of noctilucent clouds.

From his analysis of the mesospheric data obtained by the Meteorological Rocket Network (MRN), Webb (1965 and 1966) has also shown the global atmospheric circulation, indicating the coldest summer polar mesopause and ascending flow in the regions of noctilucent clouds. Based on this analysis, he presents a morphology of noctilucent clouds. His main points are as follows:

1. Size and density distribution of the noctilucent cloud nuclei (10^{-6} - to 10^{-5} - cm radius and 0.1 to 10 per cm^3, respectively) is consistent with meteoroid influx (Whipple, 1963).

2. Enhancement of meteoroid summer influx due to the forward-oriented northern hemisphere in the earth's orbit is another support for the meteorite origin of noctilucent cloud nuclei, (Fogle and Haurwitz, 1966). Note that these micrometeorites fall very fast until they hit the earth's dense atmosphere at altitudes called "meteor heights" (70-100 km – Roper, 1966) where meteors from outer space evaporate or disintegrate into micrometeorites. The terminal fall velocities of these micrometeorites below these levels decrease rapidly to the order of several 10 cm per sec.

*He also shows that the maximum frequency of noctilucent cloud occurrence is earlier at lower latitudes; i.e., it occurs in early July for $\lambda \lesssim 50°$, late July at $\lambda \simeq 60°$ and early August for $\lambda > 60°$).

3. Viewed on a global scale, most water vapor diffuses into the stratosphere at the equatorial tropopause.* This moist stratosphere air is transported toward the pole by the poleward meridional circulation that predominates in the summer hemisphere at around 50 km** altitude.

4. Because of oblique insolation at the higher latitude, the height of the region of effective heat input increases with latitude. Correspondingly, the stratopause and stratospheric circulation shift upward from the tropical atmosphere to the summer pole atmosphere.

5. Because of continuous heating such as solar ultraviolet absorption by atmospheric ozone, there is a stable anticyclonic circulation in the summer polar stratosphere.

By making a superposition of the atmospheric diurnal tide on this anticyclonic circulation, Webb (1966) has shown a strong easterly wind in the night side of the summer polar stratosphere, and called it "stratospheric tidal jet." This is shown in Figure 21. The flow pattern in this figure, however, has no theoretical foundation. It should be noted that according to the theory of the thermally driven diurnal tide of the earth's atmosphere, amplitudes of tidal motions decrease in the high latitude although they increase with altitude (Lindzen, 1967). As pointed out in section 2.3, the cold-in-summer mesopause as well as the warm-in-winter mesosphere can be regarded as the result of the thermally driven global scale annual oscillation of the earth's atmosphere (Sawada and Matsushima, 1964). The anticyclonic and upward flow in the midnight section of the summer polar mesosphere agree very well with the general southwestward movement of noctilucent clouds and the suspension of cloud-nuclei.

Chapman and Kendall (1966) develop a more rigorous mathematical theory of the noctilucent clouds, starting from the following assumptions: (1) the cloud particles consist of water vapor condensed on dust nuclei; (2) the dusts are of meteoric origin, descending slowly by diffusion under gravity until they hit the "turbopause," (the upper boundary of the turbosphere where turbulent motion predominates), then dispersing quickly toward the lower atmosphere; (3) water vapor has been brought up by convection to the mesopause; above the mesopause water vapor content decreased rapidly with altitude due to the dissociation by

*Thus, the effectiveness of the cold trap at the tropopause and consequent dry stratosphere are questioned. The tropopause cold acts only to regulate the diffusive flow and thus increase the time required to achieve equilibrium in the moist lower stratosphere (Webb, 1965).

**This height and flow velocity increase at higher latitudes. The average northward velocity is 3 m per sec at $\lambda = 20°$, but exceeds 10 m per sec at $\lambda = 60°$.

solar ultraviolet radiation; (4) the turbopause occasionally descends from its usual height of 100 km to the mesopause (around 80 km), where brought-up water vapor condenses on the nuclei, forming the cloud particles that are seen by reflected sunlight.

A mathematical discussion is outlined below.

Since the sizes of falling dust-nuclei are less than the mean free path of air molecules (which is of the order of 1 cm at 80 km) Stokes' formula cannot be used to calculate their velocities of fall. Diffusion of dust particles whose radii are of the order of 10^{-6} cm in the air is discussed by calculating the "vertical diffusing flux." The diffusing flux consists of three parts: (1) ordinary diffusion, which tends to reduce the concentration gradient and make the composition uniform; (2) gravity diffusion, by which the heavier element descends and the lighter one ascends; and (3) thermal diffusion, by which particles in the warmer region diffuse into the colder region, producing the highest concentration in the coldest part of the system.

Since temperature increases with height above the mesopause, thermal diffusion tends to cause descent of dust particles. It is shown, however, that particle descent is due predominantly to gravity.* It is interesting to note that the Stokes-Cunningham equation (which is the Stokes' equation with a correction factor attributed to "slip of gas at the surface of the dust particle,"**) gives essentially the same value of fall speed as the gravity diffusion equation. ***

Since the molecules of water vapor are lighter than average air molecules and since the atmospheric temperature decreases with altitudes below the mesopause, both gravity diffusion and thermal diffusion cause the water vapor to ascend to the mesopause. For the mesosphere with a lapse rate of -5°C per km, it is shown that the gravity-diffusion rate of water vapor is of the order of

---

*For example, in the thermosphere with temperature gradient 5° C per km at mesopause temperature 150° K, gravity-diffusion speed for dust particles of radius 2.10^{-6} cm and density 3 (specific weight) is 18 cm per sec downward whereas the thermal-diffusion speed is only 3.3×10^{-3} cm per sec downward.

**This is because the molecular mean free path is comparable with or greater than the size of the dust particle.

***Writing the radius and density of the dust particle as r and ρ, respectively, fall speed at the noctilucent cloud level is 3×10^{6} r ρ cm per sec by the gravity diffusion and 2×10^{6} r ρ cm per sec by the Stokes-Cunningham formula, respectively (Fogle, 1966).

2×10^{-2} cm per sec, (i.e., 17.3 m per day) and the thermal-diffusion rate of the order of 10^{-3} cm per sec. Because of the change of thermal gradient above the mesopause, upward diffusion of water vapor decreases. Furthermore, photodissociation by solar ultraviolet radiation increases with altitude. Therefore, as shown by Bates and Nicolet (1950), the concentration of water vapor decreases rapidly at altitudes between 80 km and 95 km, giving the scale height of the order of 2 km.

When the turbopause hits the mesopause, the dust particles are dispersed by eddy diffusion. The coefficients of eddy diffusion in these altitudes are still the subject of speculation. The assumption that $D_E = 10D$ below mesopause and zero above mesopause, (where D_E is the coefficient of eddy diffusion and D is the coefficient of molecular diffusion) implies the formation of a sharp ledge of dust at the mesopause. The results of computations done by Chapman and Kendall (1965) can be therefore summarized as follows: (1) because of the increasing photodissociation by solar ultraviolet light, the thickness of the upward-diffusing water-vapor layer is about 1.5 km; (2) because of eddy diffusion, dusts and noctilucent cloud particles are immediately dispersed below the mesopause, forming a sharp visible ledge of the clouds; (3) the upward-convecting velocity of 18 cm per sec to maintain the dusts at the mesopause can be reduced by hoar frost or flake-type water coatings of 10^{-6}-cm; and (4) noctilucent clouds can appear only when the turbopause (whose usual height is around 100 km) is lowered to the mesopause at about 80 km. The mechanism of lowering the turbopause is left open. As pointed out by Webb (1966), lowering of the mesospheric temperature to saturate water vapor might be another important factor controlling the appearance of noctilucent clouds, it still seems necessary to measure temperature distribution, turbulence, and the concentration of cloud-particles.

4.3 Noctilucent Clouds and Aurorae

By time-lapse photography, Fogle (1966) has shown not only the motion of noctilucent clouds but also the occasional appearance of aurorae and their effect on noctilucent clouds. The general effects are: (1) a reduction of the brightness and the extent of noctilucent clouds; (2) transformation of a well-ordered structure into swirls or formless veils; and (in some cases) (3) disappearance of the noctilucent clouds. These observations indicate that the precipitation of auroral particles causes turbulence in the layer of noctilucent clouds and, in an acute case, might cause even the heating of the cloud level. As discussed in section 3.4, rapid variations of auroral electron precipitation causes dynamic disturbances in the mesosphere. If these disturbances exceed a certain level, it is highly probable that the cloud particles around mesopause will be rapidly dispersed by eddy diffusion as shown by the computations of Chapman and Kendall (1965). Subsequently,

the luminosity of the cloud diminishes. For auroral electrons to cause direct heating of the cloud layer, their energy should exceed the order of 100 kev. The precipitation of these high-energy electrons is not common in the usual bright aurora. To estimate the heating of the cloud layer by precipitating auroral electrons, it seems desirable to make direct measurements of atmospheric temperature around the noctilucent cloud level. If temperature variation is negligible during the diminishing phase of noctilucent clouds following auroral activity, the coefficient of eddy diffusion around the noctilucent cloud level can be derived from its time variations.

5. Summary

Due to prolonged dark-night, twilight and daytime, there are peculiar annual variations imposed on the diurnal change in the polar atmosphere. Since the geographic poles are close to the geomagnetic poles, the effects of charged particles emitted from the sun are far stronger in the polar upper atmosphere than at other latitudes. Most of these charged particles follow the magnetic lines of force around the earth, as well as in interplanetary space. In order to understand the polar atmospheric phenomena, therefore, we must know about the magnetic fields around the earth and the sun. After the successful launching of artificial satellites and subsequent discoveries of radiation belts, our old, pre-IGY concepts of phenomena above the ionosphere have been completely changed. Furthermore, many new facts have been discovered and investigated, for instance: the deformed earth's magnetic field holding plasma (radiation belts in the magnetosphere), the outer boundaries of the magnetosphere (magnetopause), shock waves in front of the day-side magnetopause, and the geomagnetic tail in the night-side magnetopause, etc. It is impossible to describe these new environmental findings (perhaps the most exciting fields of discovery today) in this short review; they will be discussed at some future time.

To indicate the present concept of the earth's environment, Figure 22 shows one of the latest models of the earth's magnetosphere and its tail, which has been built by a series of satellite observations on the magnetic field and plasma (Ness, 1966). These models can explain complicated auroral phenomena, such as spatial and temporal variations of auroral arcs. However, the origin of auroral electrons and their acceleration mechanisms are still a great mystery. Dr. Hess of the Goddard Space Flight Center* (Hess, 1966) plans to produce an artificial aurora by ejecting high-energy electron beams from an accelerator (order of 10 kev, pulse output

*Now at the Manned Spacecraft Center, Houston, Texas.

up to 15 kw) mounted on a satellite. This will be tested in the near future, and will be a great contribution to geophysics and space science.

Finally, the importance of rocket soundings should be emphasized. Since such important events and phenomena as aurorae, PCA-airglow and noctilucent clouds, etc., are far beyond a balloon's ceiling height ($\sim$40 km) but below satellite orbits ($>$100 km), rockets are the only instruments by which these phenomena can be observed directly. (Low-orbit, short-life satellites could be developed but would be less efficient and reliable.)

ACKNOWLEDGMENT

This review paper was written at the request of the editor of "TENKI", a journal of the Meteorological Society of Japan. I am grateful for many helpful comments and criticisms from the following colleagues: Drs. A. C. Aikin, S. I-.Akasofu, S. J. Bauer, Y. Hakura, and W. Nordberg.

REFERENCES

Akasofu, S. -I., "Thickness of an Active Auroral Curtain," J. Atmos. Terr. Phys. 21:287-288, 1961.

Akasofu, S. -I., "The Development of the Auroral Substorm," Plan. Space Sci. 12:273-282, 1964.

Akasofu, S. -I., S. Chapman and C.-I. Meng, "The Polar Electrojet," J. Atmos. Terr. Phys. 27:1275-1305, 1965.

Akasofu, S.-I., S. Chapman, and A.B. Meinel, "The Aurora," Handbuch der Physik (ed. by S. Flügge), 49:1, Geophysik III, Berlin: Springer-Verlag, 1966.

Alfvén, H., "Cosmical Electrodynamics," Oxford Univ. Press, 1950.

Appleton, E.V., "Regularities and Irregularities in the Ionosphere," (The Bakerian Lecture) Proc. Roy. Soc. 162A, 451-478, 1937.

Appleton, E.V., and W.R. Piggott, Proc. Intn'l Scientific Radio Union 7:320, 1954.

Barth, C.A., Space Research V, p. 767, 1964.

Bates, D.R., and M. Nicolet, "The Photochemistry of Atmospheric Water Vapor," J. Geophys. Res. 55:301-327, 1950.

Belrose, J.S., Proc. of the NATO Advanced Study Institute, Norway, 1963.

Belrose, J.S., "Physics of the Earth's Upper Atmosphere," Prentice Hall Ins., New York, 1965.

Benyon, W.J.G., and K. Davies, "A Study of Vertical Incidence Ionospheric Absorption at 2 Mc/s," Proc. Conf. Phys. Ionosph., p.40, Special issue: Proc. Phys. Soc., 1965.

Birkeland, K., "Sur les Rayons Cathodiques sous l'action de forces magnétiques intenses," Arch. Soc., Phys. Naturelles 1:497-512, 1896.

Birkeland, K., "On the Cause of Magnetic Storms and the Origin of Terrestrial Magnetism," 1st section 1908, 2nd section 1913, in "The Norwegian Auora Polaris Expedition, 1902-1903," Aschelong, Oslo, Norway, 1908 and 1913. Vol. 1.

Bossolasco, M., and A. Elena, "Absorption de la couche D et température de la mésosphère," Compt Rend. 256:4491–4493, 1963.

Bossolasco, M., and A. Elena, "Stratospheric Warming and Ionospheric Absorption in Winter," Symposium on Interaction Between Upper and Lower Layers of the Atmosphere, Vienna, May 1966.

Chamberlain, J.W., "Physics of the Aurora and Airglow," Academic Press New York, 1961.

Chapman, S., "The Absorption and Dissociative or Ionizing Effect of Monochromatic Radiation in an Atmosphere on a Rotating Earth," Proc. Phys. Soc., 43, Part 1: pp. 26-45; and Part II: 483-501, 1931.

Chapman, S., and J. Bartels, "Geomagnetism," Oxford Univ. Press, London, 1940, p. 652.

Chapman, S., and P. C. Kendall, "Noctilucent Clouds and Thermospheric Dust; Their Diffusion and Height Distribution," Quart. J. Roy. Met. Soc. 91:115-131, 1965.

Chrzanowski, P., G. Greene, K. T. Lemon, and J. M. Young, "Traveling Pressure Waves Associated with Geomagnetic Activity," J. Geophys. Res. 66:3727-3733, 1961.

Cole, K.D., "Joule Heating of the Upper Atmosphere," Australian J. Phys. 15:223-235, 1965.

Coroniti, S.C., and R. Penndorf, "The Diurnal and Annual Variations of f_0F2 over the Polar Region," J. Geophys. Res. 64:5-18, 1959.

Davidson, G.T., "Expected Spatial Distribution of Low Energy Protons Precipitates in the Auroral Zones," J. Geophys. Res. 70:1061-1068, 1965.

Davis, T.N., "The Aurora" in "Introduction to Space Science: (ed. by W.N. Hess) New York: Gordon and Breach Science Pub., 1965, pp. 205-249.

Dieminger, W., J. Atmos. Terr. Phys. 2, 340-349, 1952. "Uber die Ursache der exzessiven Absorption in der Ionosphäre an Wintertagen.

Duncan, R.A., "Universal-time Control of the Arctic and Antarctic F-Region," J. Geophys. Res. 67:1823-1830, 1962.

Eather, R.H., "Auroral Proton Precipitation and Hydrogen Emissions," Rev. Geophys. 5, 207-285, 1967.

Eckart, H. C., "Hydrodynamics of Oceans and Atmospheres," New York: Pergamon Press, 1960.

Fogle, B., "Noctilucent Clouds in Alaska During 1962," Nature 199: 1080, 1962.

Fogle, B., "Noctilucent Clouds in the Southern Hemisphere," Nature 204:14-18, 1964.

Fogle, B., "Noctilucent Clouds over Punta Arena, Chile," Nature 207:66, 1965.

Fogle, B., "Noctilucent Clouds, "UAG R-177, Scientific Report, Geophysical Institute, Univ. of Alaska, May 1966.

Fogle B., S. Chapman, and C. Echols, "Noctilucent Clouds - A Survey with Special Reference to Recent Observations," Geophysical Institute Report UAG R-162, 1965.

Fogle, B., and B. Haurwitz, "Noctilucent Clouds," Space Sci. Rev. 6:279-340, 1966.

Frank, L.A., J. A. Van Allen and J. D. Craven, "Large Diurnal Variations of Geomagnetically Trapped and of Precipitated Electrons Observed at Low Altitudes," J. G. R., 69, 3155-3167, 1964. to p. 26.

Fukushima, N., "Polar Magnetic Storms and Geomagnetic Bays," J. Fac. Sci. Univ. Tokyo, Section II, Part V, 8, 291-412, 1953.

Gregory, J. B., "Radio Wave Reflections from the Ionosphere, 1. Heights of Occurrence," J. Geophys. Res. 66:429-445, 1961.

Gregory, J. B., "The Influence of Atmospheric Circulation on Mesospheric Electron Densities in Winter," J. Atmos. Sci., 22:18-23, 1965.

Gregory, J.B., "Evidence for Penetration of Mobile Waves from the Troposphere to the Mesosphere," Symposium at Vienna, May 1966.

Harang, L., "The Aurorae," New York: John Wiley and Sons, 1951.

Haurwitz, B., "Comment on Wave Forms in Noctilucent Clouds," Geophys. Inst. Report No. UAG-R-160, Univ. of Alaska, 1964.

Heppner, J.P., "Study of Relationiships between the Aurora Borealis and the Geomagnetic Disturbances Caused by Electric Currents in the Ionosphere," Report No. DR-135, Thesis Calif., Inst. Tech., Pasadena, 1958.

Hess, W.N., "Electron Beam Field Mapping," NASA Goddard Space Flight Center Document X-640-65-492, 1966.

Hill, G.E., "Anomalous to F_2 variations in the Antarctic," J. Geophys. Res. 65:2011-2023, 1960.

Hines, C.O., "Internal Gravity Waves at Ionospheric Heights," Canad. J. Phys. 38:144-1481, 1960.

Jesse, O., "Die Höhe der Leuchtenden Nachtwolken," Astr. Nachtr 140:161, 1896.

Johnson, F.A., "The Physical Properties of the Earth's Ionosphere," (preprint) 1960.

Jones, L. M., J. W. Peterson, E. J. Schaefer, and H. F. Schultz, "Upper-Air Density and Temperature; Some Variations and an Abrupt Warming in the Mesosphere," J. Geophys. Res. 64:2331-2340, 1959.

Kellogg, W. W. and G. F. Schilling, J. Met. 8, 222, 1951.

Kellogg, W. W., "Chemical Heating above the Polar Mesopause in Winter," J. Met. 18, 371-381, 1961.

Knecht, R. W., Progress in Radio Science, Vol. 3, 1965.

Lindzen, R. S., "Thermally driven diurnal tide in the atmosphere," Quart. J. Roy. Met. Soc. 93: 18-42, 1967.

Maeda, K., and T. Watanabe, "Pulsating Aurorae and Infrasonic Waves in the Polar Atmosphere," J. Atmos, Sci. 21:15-29, 1964.

Maeda, K., and J. Young, "Propagation of the Pressure Waves Produced by Auroras," J. Geomag. Geocl. 18:275-299, 1966.

Maehlum, B., "On the 'Winter Anomaly' in the Mid-Latitude D-Region," J. Geophys. Res. 72:2287-2299, 1967.

McIlwain, C.E., "Coordinates for Mapping the Distribution of Magnetically Trapped Particles." J. G. R., 66, 3681-3691, 1961.

Meinel, A. B., "Doppler-Shifted Auroral Hydrogen Emission," Astrophys. J. 113:50-54, 1951.

Murgatroyd, R. J., and F. Singleton, "Possible Meridional Circulations in the Stratosphere and Mesosphere," Quart. J. Roy. Met. Soc. 87:125-135, 1961.

Murgatroyd, R. J., "Winds and temperatures between 20 km and 100 km – a review," Quart. J. Met. Roy. Soc. 38:417-458, 1957.

Nagata, T. and N. Fukushima," Constitution of Polar Magnetic Storms," Rep. Iono. Res. Japan 6, 75-97, 1952.

Narcisi, R. S., and A. D. Bailey, "Mass Spectrometric Measurements of Positive Ions at Altitudes from 64 to 112 Kilometers," J. Geophys. Res. 70:3687-3700, 1965.

Ness, N. F., Preprint presented at the Belgrade Symposium, 1966.

Nicolet, M., "Nitrogen Oxides in the Chemosphere," J. Geophys. Res. 70:679-689, 1965.

Nicolet, M., and A.C. Aikin, "The Formation of the D-Region of the Ionosphere," J. Geophys. Res. 65:1469-1483, 1960.

Nordberg, W., L. Katchen, J. Theon, and W. S. Smith, "Rocket Observations of the Structure of the Mesosphere," J. Atmos. Sci. 22:611-622, 1965.

Paton, J., "Luminous Night Clouds," Met. Mag. 78:354- 1949.

Paton, J., "Aurora and Luminous Night Clouds," Proc. Phys. Soc. Lond. 63B:1039, 1950.

Paton, J., "Direct Evidence of Vertical Motion in the Atmosphere at a Height of About 80 km Provided by Photographs of Noctilucent Clouds," Proc. Toronto Met. Conference 1953, 31-33, 1954.

Paton, J., "Noctilucent Clouds," Met Mag. 93:161-179, 1964.

Paton, J., "Noctilucent Clouds in 1964," Met. Mag. 94:180-184, 1965.

Piddington, J.H., "Geomagnetic Storms, Auroras and Associated Effects," Space Science Rev. 3:724-780, 1964.

Ratcliffe, J.A., "The Magneto-Ionic Theory and its Application to the Ionosphere," Cambridge Univ. Press, London 1959.

Ratcliffe, J. A., "Physics of the Upper Atmosphere," Academic Press, New York 1960, p. 429.

Roper, R. G., "Atmospheric Turbulence in the Meteor Region," J. Geophys. Res. 71:5785-5792, 1966.

Sanford, B. P., "Polar-Glow Aurora in Polar Cap Absorption Events," J. Atmos. Terr. Phys. 24: 155-171, 1962.

Sato, T., and G. F. Rourke, "F-Region Enhancements in the Antarctic," J. Geophys. Res. 69:4591-4607, 1964.

Sawada, R. and A. Matsushima," Thermally Driven Annual Atmospheric Oscillations as a Cause of Dynamic Heating of the winter Polar Upper Mesosphere," J. Met. Soc., Japan 42, 97-108, 1964.

Sechrists, C. F., Jr., "A Theory of the Winter Absorption Anomaly, at Middle Latitudes," J. Atmos. and Terr. Phys. 29:113-136, 1967.

Shapley, A. H., and W. J. G. Baynon, Nature 206:1342, 1965.

Singer, S. F., "Zodiacal Dust and Deep-Sea Sediments," Science 156:1080-1083, 1967.

Soberman, R. K., "Noctilucent Clouds," Sci. Amer. 208:50-59, June 1963.

Spitzer, L., "Physics of Fully Ionized Gases," Interscience Pub. N.Y. 1956.

Störmer, C., "Measurements of Luminous Night Clouds in Norway 1933 and 1934," Astrophysica Norwegica 1:87-114, 1935.

Störmer, C., "The Polar Aurora," Oxford: Clarendon Press, 1955.

Stroud, W.G., W. Nordberg, W.R. Bandeen, F. L. Bartman, and P. Titus, "Rocket-Grenade Observations of Atmospheric Heating in the Arctic," J. Geophys. Res. 64:1342-1344, 1959.

Stroud, W. G., W. Nordberg, and W. R. Bandeen, "Rocket Grenade Measurements of Temperatures and Winds in the Mesospheric over Churchill, Canada," J. Geophys. Res. 65:2307-2323, 1960.

Theon, J. S., W. Nordberg, and W. S. Smith, "Temperature Measurements in Noctilucent Clouds," Science 157:419-421, 1967.

Thomas, L., "Radiowave Absorption in the Ionosphere," Pergamon Press, New York 301, 1962

Vestine, E. H., Trans. Wash. Assem. Publ. IA TME Bull. No. 11, 360, 1940.

Vestine, E. H., "Noctilucent Clouds," J. Roy. Astron. Soc. Can. 28:249-272 and 303-317, 1934.

Webb, W. L., "Morphology of Noctilucent Clouds," J. Geophys Res. 70:4463-4475, 1965.

Webb, W. L., "Structure of the Stratosphere and Mesosphere," New York: Academic Press, 1966 a, p. 382.

Webb, W. L., "The Stratospheric Tidal Jet," J. Atmos. Sci. 23:531-534, 1966 b.

Webb, W. L., "Stratospheric Tidal Circulations," Rev. Geophys: 4:363-375, 1966 c.

Whipple, F. L., "On Meteoroids and Penetration, "J. Geophys. Res. 68:4929-4939, 1963.

Wilson, C. R., and S. Nichparenko, "Infrasonic Waves and Auroral Activity," Nature 214-1299-1302, 1967.

Witt, G., "Polarization of Light from Noctilucent Clouds," J. Geophys. Res. 65:925-933, 1960 a.

Witt, G., "A Note to the Paper by G. Witt, Polarization of Light from Noctilucent Clouds," J. Geophys. Res. 65:2199-2200, 1960b.

Witt, G., C. L. Hemenway, and R. K. Soberman, "Collection and Analysis of Particles from the Mesopause," Space Research IV:197-203, 1964.

Witt, G., C. L. Hemenway, N. Lange, S. Modin, and R. K. Soberman, "Composition Analysis of Particles from Noctilucent Clouds," Tellus 16:103-109, 1964.

APPENDIX A

The Ionospheric Critical Frequency (Plasma Frequency)

The wave equation for the electromagnetic wave is derived from Maxwell's equations for Faraday's and Ampere's laws by eliminating magnetic field $\vec{H}$. In vacuum this is

$$\nabla^2\vec{E} = \frac{1}{c^2}\frac{\partial^2\vec{E}}{\partial t^2}, \tag{A.1}$$

where $\vec{E}$ is the electric field intensity, c is the speed of light in vacuum.

In plasma, because of the mobility of ions and electrons which cause an electric current $\vec{j}$ in the plasma, the wave equation is (instead of Equation A.1):

$$\nabla^2\vec{E} = \frac{1}{c^2}\frac{\partial^2\vec{E}}{\partial t^2} + \frac{4\pi Z\rho e^2}{m_i m_e c^2}\vec{E} \qquad \text{(Spitzer, 1956).} \tag{A.2}$$

where $\rho = n_i m_i + n_e m_e$, n_i and m_i are number density and mass, respectively of ions, n_e and m_e are number density and mass, respectively, of electrons, Z is charge of ion.

Substitution of $\vec{E} = \vec{E}_0 e^{i(\vec{k}\cdot\vec{r} - \omega t)}$ in Equation (A.2) results in the following dispersion relation:

$$-k^2 = -\frac{\omega^2}{c^2} + \frac{\omega_p^2}{c^2} \tag{A.3}$$

where

$$\omega_p^2 = \frac{4\pi Ze^2 (n_i m_i + n_e m_e)}{m_i m_e} \tag{A.4}$$

$$= \frac{4\pi n_e e^2}{m_e}\left(1 + Z\frac{m_e}{m_i}\right).$$

In the second expression of Equation A.4, the neutrality of plasma is assumed (i.e., $Zn_i = n_e$). From Equation A.3 the phase velocity of the wave, $\bar{V}$, in the plasma is given by

$$\bar{V} = \frac{\omega}{k} = \frac{c}{\left(1 - \frac{\omega_p^2}{\omega^2}\right)^{\frac{1}{2}}} . \tag{A.5}$$

The index of refraction, n, of the plasma is therefore

$$n = \frac{c}{\bar{V}} = \sqrt{1 - \frac{\omega_p^2}{\omega^2}} . \tag{A.6}$$

By Equation A.6, the index of refraction is imaginary if $\omega < \omega_p$. In other words, a wave whose frequency is lower than the local plasma frequency cannot propagate the medium. Since $m_e << m_i$, the second term in Equation A.4 can be neglected, and, noting $f_p = \omega_p / 2\pi$, Equation A.4 gives

$$f_p = \left(\frac{e^2}{m_e} \cdot \frac{n_e}{\pi}\right)^{\frac{1}{2}} \simeq 9 \cdot 10^3 \sqrt{n_e} . \tag{A.7}$$

This leads to Equation 2.1 for the F2-peak critical frequency. Generally f_p, given by Equation A.7 is called "plasma frequency"* and is a function of electron concentration only.

---

*ω_p should be called plasma angular frequency.

APPENDIX B

The Chapman Layer

The variation of electron density, N, at a fixed altitude, z , is given by

$$\frac{dN}{dt} = q - L, \tag{B.1}$$

where q and L are the rates of electron production and loss, respectively.

Loss L in Equation B.1 consists of two processes: recombination of electrons with positive ions and attachment to neutral molecules. Both losses have rates proportional to N^2 , i.e.,

$$L = \alpha N^2, \tag{B.2}$$

where α is a constant.

The production rate, q , due to photoionization of air molecules by the solar ultraviolet radiation, is proportional to the radiation intensity I (z), which can be given as a solution of the attenuation equation (Chapman, 1931),

$$\frac{dI}{I} = -A \cdot \rho(z) \cdot \sec\chi \cdot dz, \tag{B.3}$$

where A = atmospheric absorption coefficient of radiation.

$\rho(z)$ = air density at altitude z, and

χ = sun zenith angle of the sun.

Assuming that $\rho(z) = \rho_0 \cdot e^{-z/H}$, where H is the scale height and $\rho_0 = \rho(z = 0)$, the solution of Equation B.3 is

$$I(z) = I_\infty \exp\{-A\rho_0 H \cdot (\sec\chi) \cdot e^{-z/H}\}, \tag{B.4}$$

where I_∞ is the intensity of solar radiation outside the earth's atmosphere. If the number of ions produced by the absorption of unit quantity of the radiation is γ, the rate of production of ions per unit volume is

$$q(z) = \gamma \cos\chi \frac{dI}{dz}$$

$$= AI_\infty \rho_0 \exp\left\{ -\frac{z}{H} - A\rho_0 H(\sec\chi) e^{-z/H} \right\} \qquad \text{(B.5)}$$

Denoting the height of the maximum production rate, by z_m, which is obtained by $dq/dz = 0$, the rate of maximum ion production $q_m = q(z_m)$ is given by

$$q_m = \frac{\gamma I_\infty}{H \exp(1)} \cos\chi \qquad \text{(B.6)}$$

where

$$\exp(1) \cong 2.71828 \ldots *$$

The maximum number density of electrons at equilibrium is given by Equation B.1, with $dN/dt = 0$ and by Equation B.2, i.e., $q_m = \alpha N^2$.

*That is, the base of natural logarithms; its usual symbol, "e," cannot be used here because "e" is in use as the electronic charge.

Combining this with Equation B.6 gives

$$N = \left\{ \frac{\gamma}{\alpha} \cdot \frac{I_\infty}{H \exp(1)} \cdot \cos \chi \right\}^{\frac{1}{2}} \propto \cos^{\frac{1}{2}} \chi \qquad \text{(B.7)*}$$

Substituting this in Equation A.7 gives (note $N = n_e$)

$$f_p \propto (\cos \chi)^{\frac{1}{4}}. \qquad \text{(B.8)**}$$

Assuming equilibrium at all altitudes gives a vertical distribution of electron density, which is proportional to $[q(z)]^{1/2}$, where $q(z)$ is given by Equation B.5. This is called an "equilibrium Chapman layer."

**If electron loss rate L is βN, instead of being proportional to N^2 (as in Equation B.2), the following relation takes the place of Equation B.7:

$$N = \frac{\gamma}{\beta} \cdot \frac{I_\infty}{H \exp(1)} \cdot \cos \chi, \qquad \text{(B.7')}$$

Instead of Equation B.8:

$$f_p \propto \sqrt{\cos \chi}. \qquad \text{(B.8')}$$

These equations constitute another type of "Chapman layer," useful in studying photo-chemical reactions in the upper atmosphere. The two types are called "Chapman α-layer and Chapman β-layer" respectively (Ratcliffe, 1960).

APPENDIX C

Ionospheric Absorptions

The forced movements of an electron in the upper atmosphere due to the oscillating field of an electromagnetic wave can be described by the following equation of motion:

$$m\ddot{\vec{r}} = e\vec{E} + e\dot{\vec{r}} \times \vec{B}_0 - m_e \nu \dot{\vec{r}}, \tag{C.1}$$

where $\vec{r}$, $\vec{E}$, and $\vec{B}_0$ are vectors indicating displacement of an electron, the electric field of the wave, and the static magnetic field of the earth, respectively; m_e and e are the mass and charge of an electron, respectively; and ν is the average collision frequency between an electron and heavy particles in the medium (the ionosphere), called "effective collision frequency." The last term in Equation C.1 causes electromagnetic-wave attenuation by damping the motion of electrons in the medium. Displacing an electron is equivalent to introducing an electric dipole of moment $e\vec{r}$ into the medium. The electric dipole polarization, $\vec{P}$, of the medium is therefore:

$$\vec{P} = Ne\vec{r}, \tag{C.2}$$

where N is the number density of electrons in the medium.

Since $\vec{r} \sim e^{i\omega t}$, ($\omega$ is the angular frequency of the propagating electromagnetic wave) multiplying Equation C.1 by Ne gives

$$-\omega^2 m_e \vec{P} = Ne^2\vec{E} + i\omega e\vec{P} \times \vec{B}_0 - i\omega m_e \nu \vec{P} \tag{C.3}$$

Following the standard notations for magnetroionic theory (Ratcliffe, 1959), this equation becomes

$$-\frac{\epsilon_0}{4\pi} X\vec{E} = \vec{P}(1 - iZ) - i\vec{Y} \times \vec{P}, \tag{C.4}*$$

*In Gaussian units, the electric permittivity of free space ϵ_0 is unity.

where

$$X = \frac{4\pi N e^2}{\epsilon_0 m_e \omega^2} \left(= \frac{\omega_p^2}{\omega^2} \right)$$

(cf. Equation A.4)

$$\vec{Y} = -\frac{e_0 \vec{B}_0}{m_e \omega} = -\frac{\omega_L}{\omega}, \qquad \omega_L = \frac{eB_0}{m_e}$$

and

$$Z = \frac{\nu}{\omega} .$$

Taking the Cartesian coordinates x, y, z (where $z \| \vec{k}$, $\vec{k}$ is propagation vector or direction of wave normal), $y \| \vec{B}_0$, and writing $Y_L = Y \cos\theta$, $Y_T = Y \sin\theta$, θ is the angle between $\vec{k}$ and $\vec{B}_0$) gives Equation C.4 as

$$-\frac{\epsilon_0}{4\pi} X \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} = \begin{pmatrix} 1 - iZ & iY_L & 0 \\ -iY_L & 1 - iZ & iY_T \\ 0 & -iY_T & 1 - iZ \end{pmatrix} \cdot \begin{pmatrix} P_x \\ P_y \\ P_z \end{pmatrix} . \qquad \text{(C.5)}$$

Since all components of the wave fields vary as e^{ikz} (a plane wave propagating along the z-axis), Maxwell's third and fourth equations

$$\text{curl}\, \vec{H} = \frac{\partial \vec{D}}{\partial t}, \quad \text{curl}\, \vec{E} = -\mu_0 \frac{\partial \vec{H}}{\partial t}$$

can be written as follows (noting that $\vec{D} = \epsilon_0 \vec{E} + 4\pi \vec{P}$, where μ_0; magnetic permittivity of free space):

$$\begin{aligned} -ikH_x &= i\omega(\epsilon_0 E_y + 4\pi P_y) , \\ ikH_y &= i\omega(\epsilon_0 E_x + 4\pi P_x) , \\ 0 &= i\omega(\epsilon_0 E_z + 4\pi P_z) , \end{aligned} \qquad \text{(C.6)}$$

and

$$-i\omega\mu_0 H_x = ikE_y ,$$

$$i\omega\mu_0 H_y = ikE_x , \qquad (C.7)$$

$$i\omega\mu_0 H_z = 0 .$$

Equations C.6 and C.7 give

$$\frac{k^2}{\epsilon_0\mu_0\omega^2} = 1 + \frac{4\pi}{\epsilon_0}\cdot\frac{P_x}{E_x} \qquad (C.8)$$

$$= 1 + \frac{4\pi}{\epsilon_0}\cdot\frac{P_y}{E_y} , \qquad \frac{P_x}{E_x} = \frac{P_y}{E_y}$$

Since the phase velocity of the wave is $\bar{V} = \omega/k = c/n$, and because $c^2 = \epsilon_0\,\mu_0$, Equation C.8 becomes

$$n^2 = 1 + \frac{4\pi}{\epsilon_0}\cdot\frac{P_x}{E_x} . \qquad (C.9)$$

From Equation C.5

$$\frac{P_x}{E_x} \quad \text{or} \quad \left(\frac{P_y}{E_y}\right)$$

can be eliminated in Equation C.9, giving

$$n^2 = 1 - \frac{X}{1 - iZ - iY_L R} \qquad (C.10)$$

where

$$R = \frac{E_x}{E_y} = \frac{P_x}{P_y} = \frac{-i}{Y_L}\left\{\frac{\frac{1}{2}Y_T^2}{1 - iZ - X} \pm \left(\frac{\frac{1}{4}Y_T^4}{(1 - iZ - X)^2} + Y_L\right)^{\frac{1}{2}}\right\} . \qquad (C.11)$$

Since the index of refraction n is complex, this can be written

$$n = \mu - i\eta . \qquad (C.12)$$

Then, the wave attenuation is $e^{-\eta}$ after propagating a distance $z = c/\omega = \lambda/2\pi$ with velocity c/μ. In other words, the absorption co-efficient κ, defined by

$$E = E_0 e^{-\kappa z} \tag{C.13}$$

is given by

$$\kappa = \frac{\omega}{c} \cdot \eta. \tag{C.14}$$

Combining Equations C.10 and C.12 gives

$$\kappa = \frac{1}{2c\mu} \cdot \frac{4\pi N e^2}{e_0 m_e} \cdot \frac{\nu}{(\omega \pm |\omega_L|)^2 + \nu^2} \tag{C.15}$$

$$= \frac{1}{2c\mu} \cdot \frac{\omega_p^2 \nu}{(\omega \pm |\omega_L|)^2 + \nu^2}$$

Another expression is

$$\kappa = \frac{\nu}{2c} \cdot \frac{1}{(1 \pm |Y_L|)} \cdot \left(\frac{1 - \mu^2 + \eta^2}{\mu} \right) \tag{C.16}$$

At places where $\mu >> 1$, Equation C.16 can be written, assuming $\eta^2 << 1 - \mu^2$,

$$\kappa \simeq \frac{\nu}{2c} \cdot \frac{1}{(1 \pm |Y_L|)} \cdot \left(\frac{1}{\mu} - \mu \right). \tag{C.17}$$

In the above expressions, the plus sign represents the ordinary wave; the minus sign represents the extraordinary wave.

In most of the ionosphere, it can be assumed that $\nu^2 << (\omega + |\omega_L|)^2$ and, if $\omega >> \omega_L$, (i.e., $Y << 1$), $\mu \simeq 1$ then

$$\kappa \propto \frac{N\nu}{(\omega \pm |\omega_L|)^2} \tag{C.18}$$

The reflection coefficient, ρ_R, can be given approximately as

$$-\log \rho_R \sim \frac{1}{(\omega + |\omega_L|)^2} \int_0^{h_1} N\nu \, dz \quad , \tag{C.19}$$

where $N = N(z)$, $\nu = \nu(z)$, and h_1 is the height of total reflection of the wave with angular frequency ω.

An absorption obeying this frequency law (i.e., Equations C.18 and C.19) is called a "non-deviative" absorption.

On the other hand, when $Y_L < 1$, Equation C.17 can be written:

$$\kappa = \frac{\nu}{2c}\left(\frac{1}{\mu} - \mu\right) , \tag{C.20}$$

which gives, instead of Equation C.19,

$$-\log \rho_R = \frac{1}{c}\int_0^{h_1} \nu\left(\frac{1}{\mu} - \mu\right) dz \tag{C.21}$$

$$\simeq \frac{\nu}{c}(h_g - h_f) ,$$

where $$\int \frac{dz}{\mu} = h_g \text{ and } \int \mu dz = h_f$$

are called equivalent (or group) height, and optical (or phase) height, respectively. Absorptions expressed by these terms are called "deviative" absorptions. It should be noticed that in the case of non-deviative absorption $\mu \simeq 1$; therefore $h_g \simeq h_f$ and $\kappa \simeq 0$. Deviative absorptions are negligible; only non-deviative absorption is important in frequency range $\mu \simeq 1$.

APPENDIX D

The Units of Auroral Brightness

Conversion of wave length λ (in Å) of the light emitted in the transition between two energy states E_i, E_f (in eV) of atoms or molecules is obtained from the Planck's photo quantum formula

$$h\nu = \Delta E, \tag{D.1}$$

where $h = 6.625 \times 10^{27}$ erg sec ($\simeq 4.15 \times 10^{-15}$ eV sec),

$$\nu = \frac{2\pi}{\lambda} \cdot c \quad (c \text{ is light velocity, } c \simeq 3 \times 10^{10} \text{ cm/sec})$$

and $\Delta E = E_i - E_f$.

Equation D.1 is therefore written

$$\lambda \Delta E = 1.24 \times 10^4 \text{ (eV Å)} \tag{D.2}$$

For example, the energy of primary electrons ΔE required to excite oxygen atoms to the state that emits red light, 6300 Å (see Type A aurora) is, from Equation D.2,

$$\Delta E \simeq 2 \text{ eV} \cdot$$

This is indicated in the energy diagram in Figure 11(a).

The brightness of aurorae and general airglows are generally expressed by the units called Rayleigh (R), defined by 10^6 photons per cm^2 per sec. This unit is based on the total emitted photons in the column along the line of sight toward the source per unit solid angle (steradian) at the observer, multiplied by 4π (i.e., assumed isotropic luminosity

$$4\pi J = \int_0^\infty F(r)\, dr, \tag{D.3}$$

where F(r) is rate of emission at distance r from the observer (photon/cm^3 sec); the value of $4\pi J$ is "one Rayleigh" if it is 10^6 in photons per cm^2 per sec. In other words (Chamberlain, 1961), 1 R = an apparent emission rate of one megaphoton per cm^2 per second.

As can be seen from Planck's formula (D.1), conversion to energy flux from Rayleigh units differs for different wave lengths of emitted light. Unless otherwise specified, the brightness of aurora is based on the intensity of oxygen line 5577 Å, which has been used to define the International Brightness Coefficient (IBC): IBC I, II, III, and IV corresponds to brightness (10^3 R, 10^4 R, 10^5 R, and 10^6 R respectively.

The energy flux of light emitted by the IBC-I (i.e., weak) aurora is, by the quantum Equation D.1, $3.6 \times 10^{-3} \simeq 4 \times 10^{-3}$ ergs/cm^2sec.

Table D.1
Energy Fluxes Around the Earth

Source	Energy flux (ergs cm^{-2} sec^{-1})
Sun	1.4×10^6 (solar constant)
Full moon	3 *
Total starlight	1.8×10^{-3}
Airglow (visible)	1.5×10^{-2}
OH-airglow (infrared)	2×10^{-2}
Lyman alpha (ultraviolet)	10^{-2} **
Cosmic rays	3.8×10^{-3}

*Comparable with the brightest (IBC-IV) aurora.
**The solar Lyman Alpha emission (approximately 5 ergs cm^{-2} sec^{-1}) is excluded; because of strong absorption by air, this is not observable from the ground.

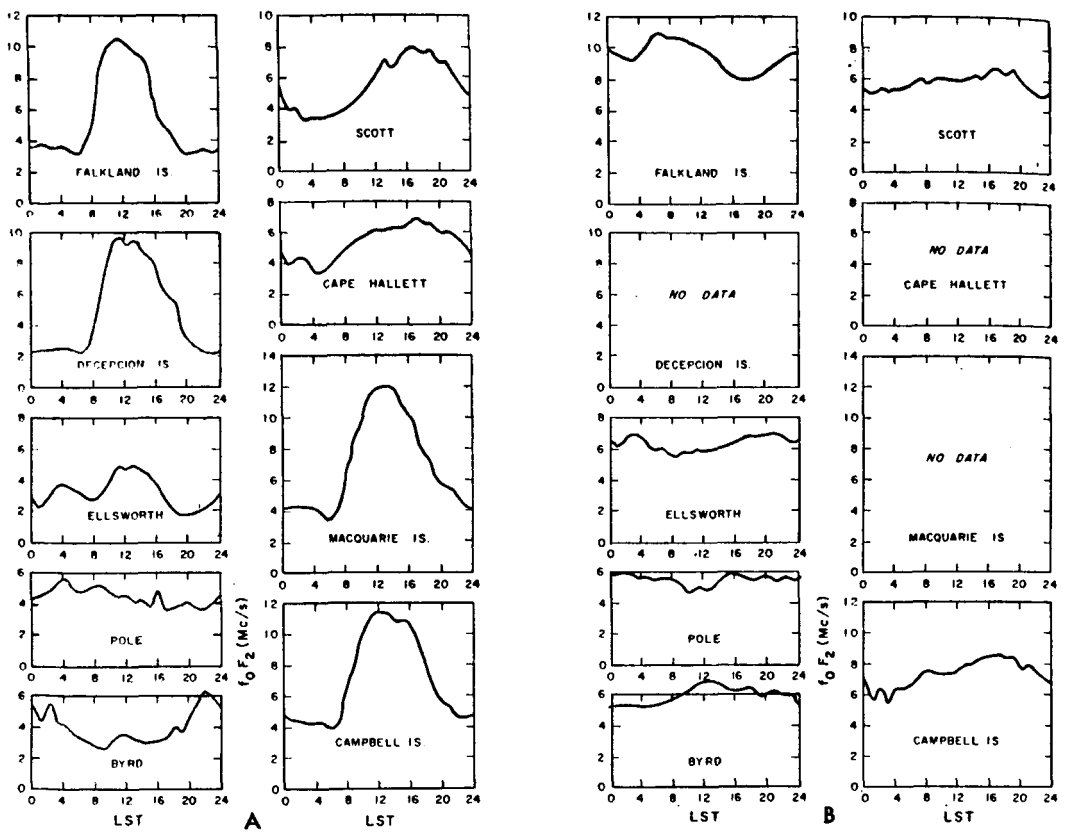
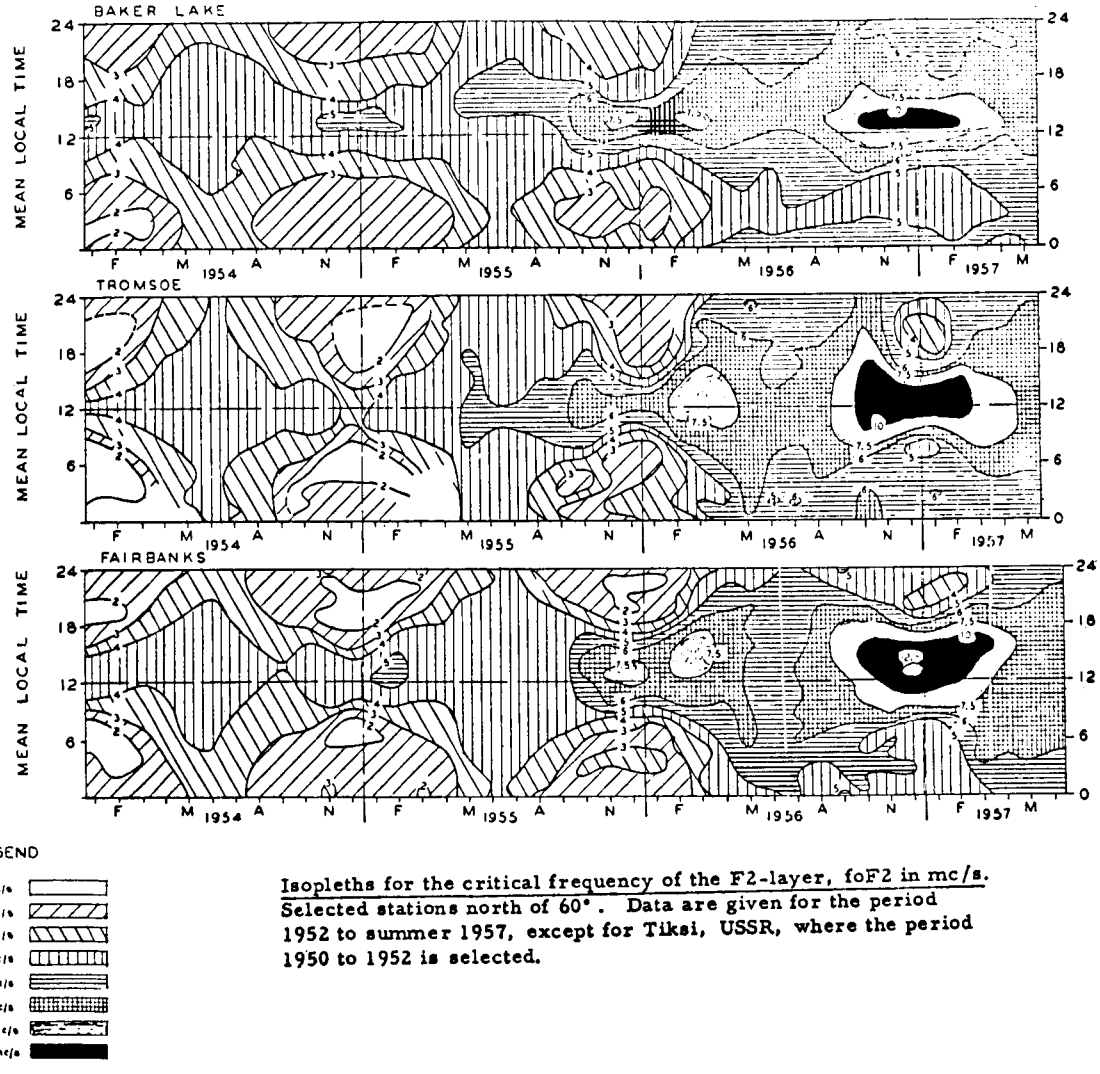


Figure 1–Diurnal variations of f_0F2 in the Antarctica based on the monthly medians of hourly values (December, 1957). (Hill, 1960).



Isopleths for the critical frequency of the F2-layer, foF2 in mc/s. Selected stations north of 60°. Data are given for the period 1952 to summer 1957, except for Tiksi, USSR, where the period 1950 to 1952 is selected.

Figure 2–Isopleths for the critical frequency of the F2-layer, f_0F2, for three-stations north of 60°N Baker Lake(64°20'N. 96°02'W: geomag. 73°8N), Tromsö (69°40'N, 18°57'E; geo-67°0N) and Fairbansk (64°51'N, 147°50'W; geomag. 67°7N), (Coroniti and Penndorf, 1959).

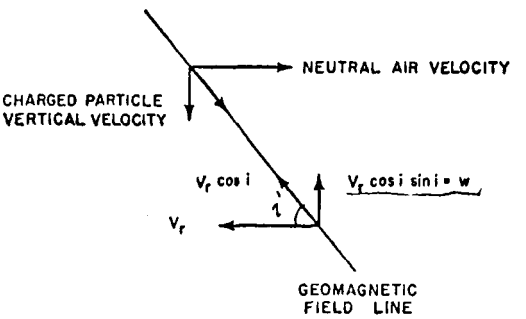


Figure 3–Effects of wind shear and geomagnetic field in the F-region.

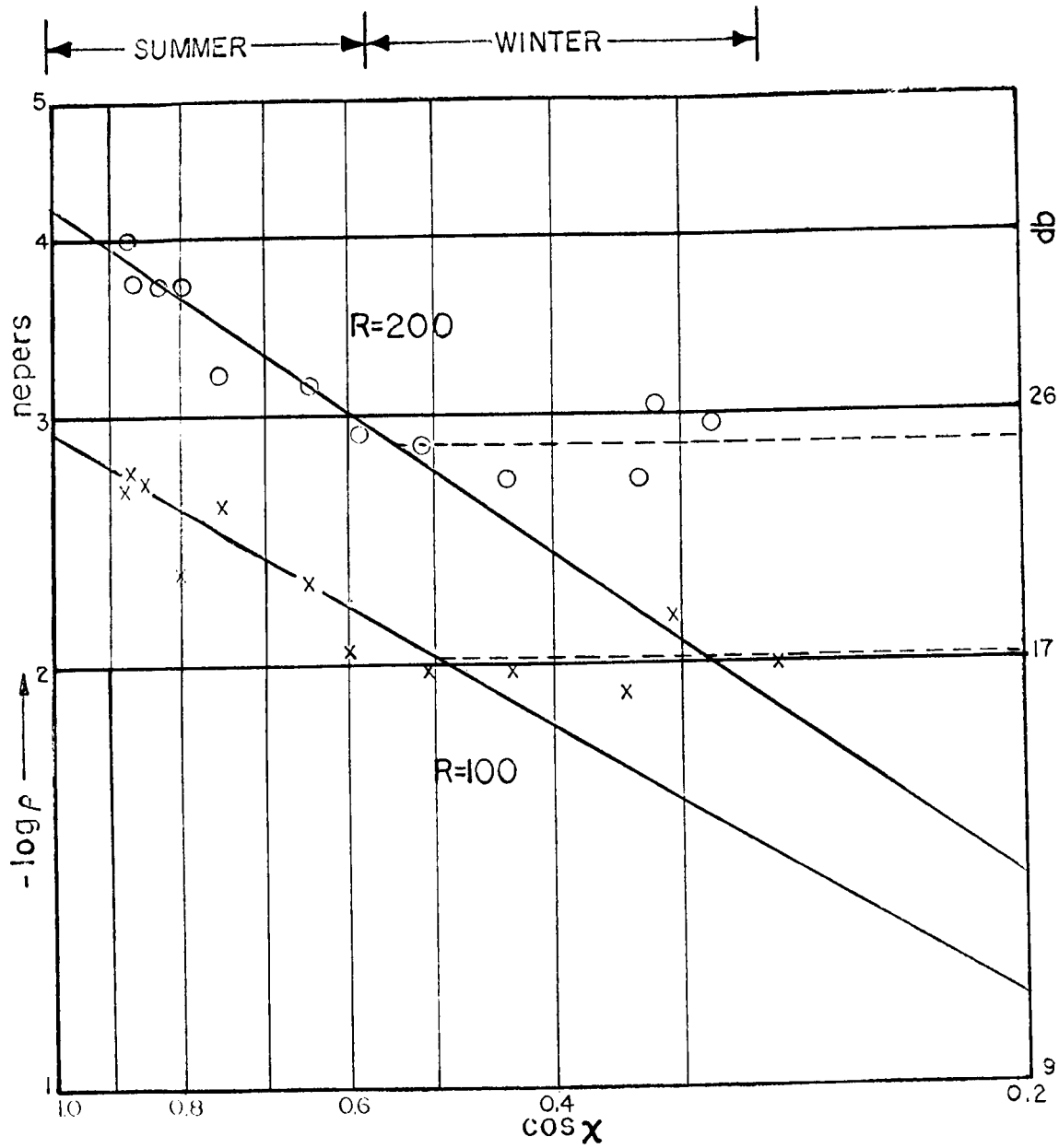


Figure 4–Average monthly values of absorption versus inverse Chapman's function $[ch\ (z,\chi)] \simeq \cos\ \chi$, for constant sunspot numbered R = 100 and 200, respectively (f = 4 MHz).

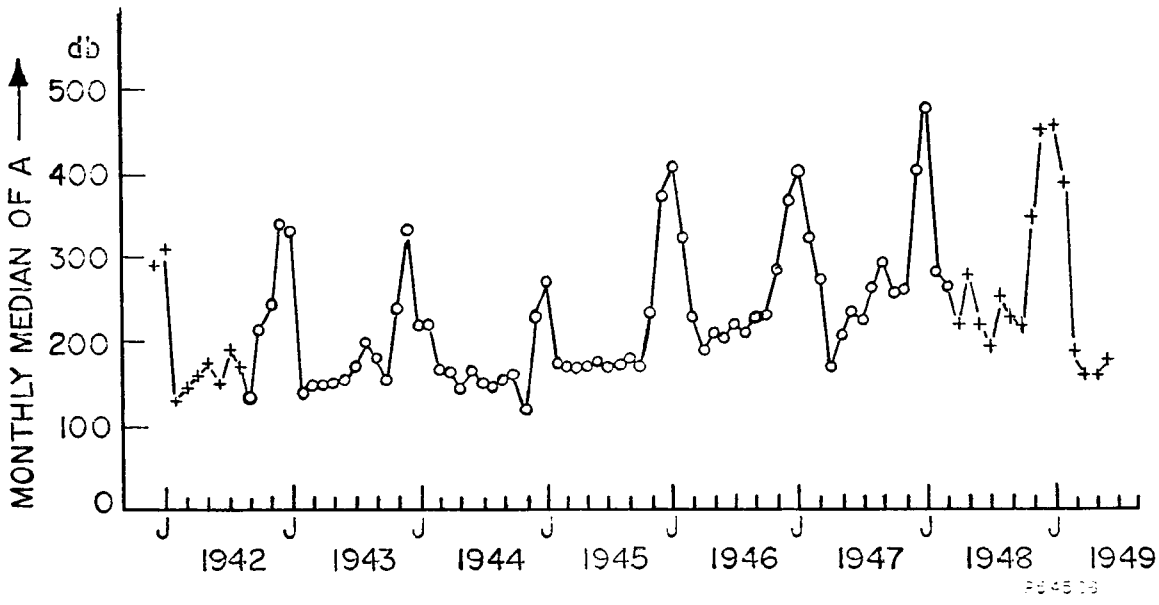


Figure 5–Variation of absorption for constant zenieth angle ($\chi \simeq 75°$) at Slough (after Appleton and Piggett, 1954).

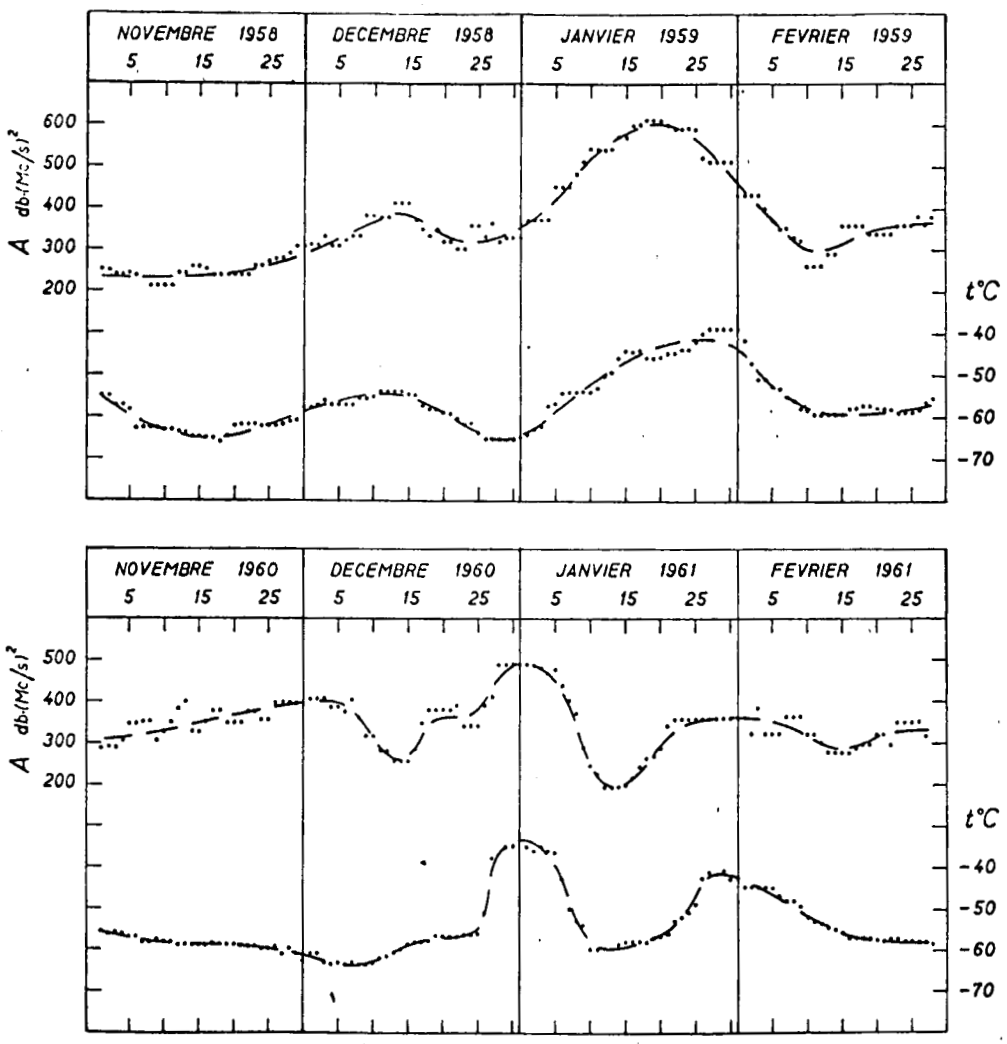


Figure 6–Temporal variations of the stratospheric temperature and non-deviative absorption in Central Europe. (Bossolasco and Elena, 1963).

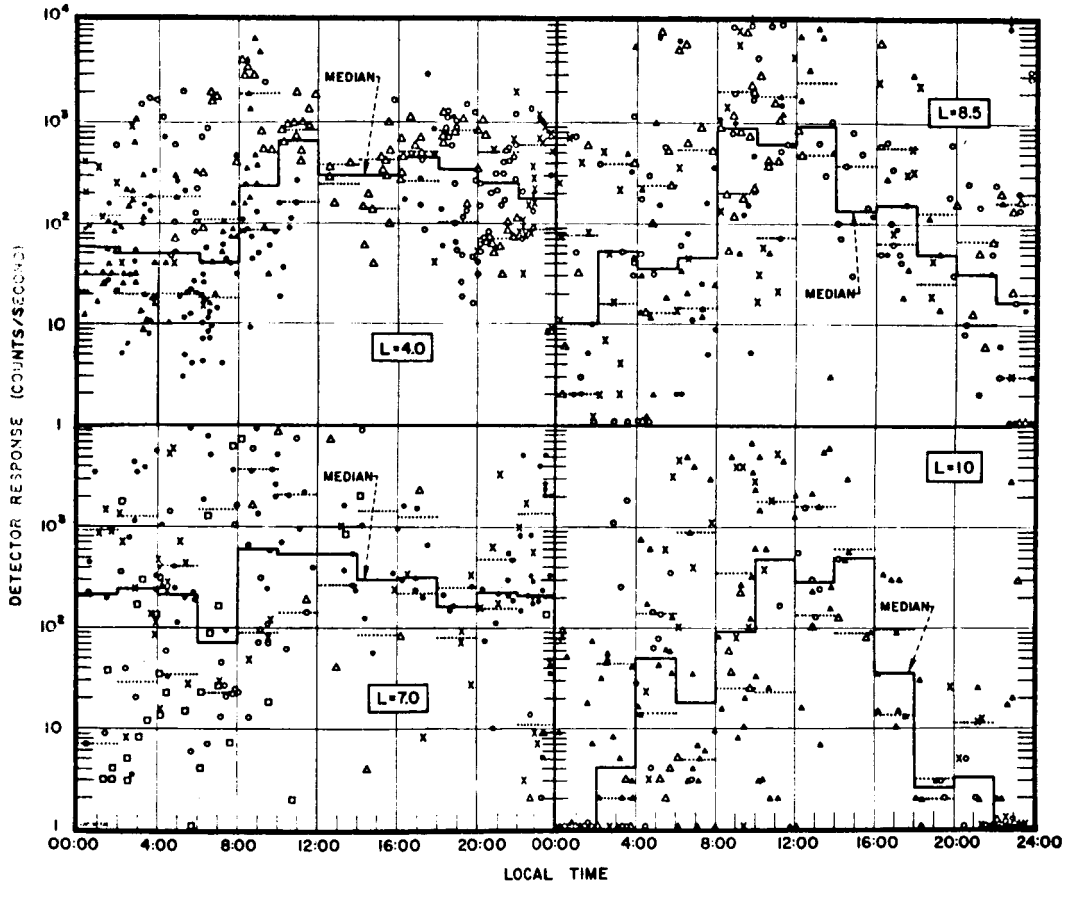


Figure 7–Local time-dependence of electron precipitation (E > 40 kev) by satellites Injun 3 (Frank, Van Allen, Craven, 1964).

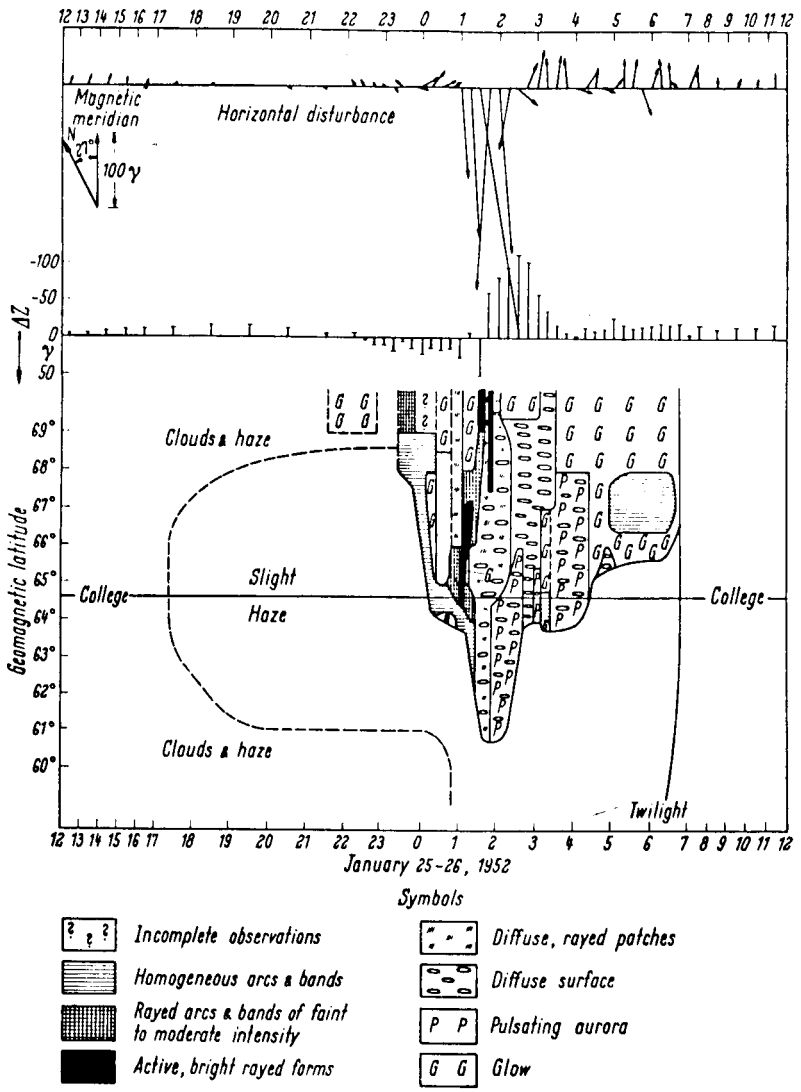


Figure 8–Temporal and spatial (horizontal) distribution of the auroral display observed at College, Alaska, on January 25-26, 1952 (Heppner, 1958).

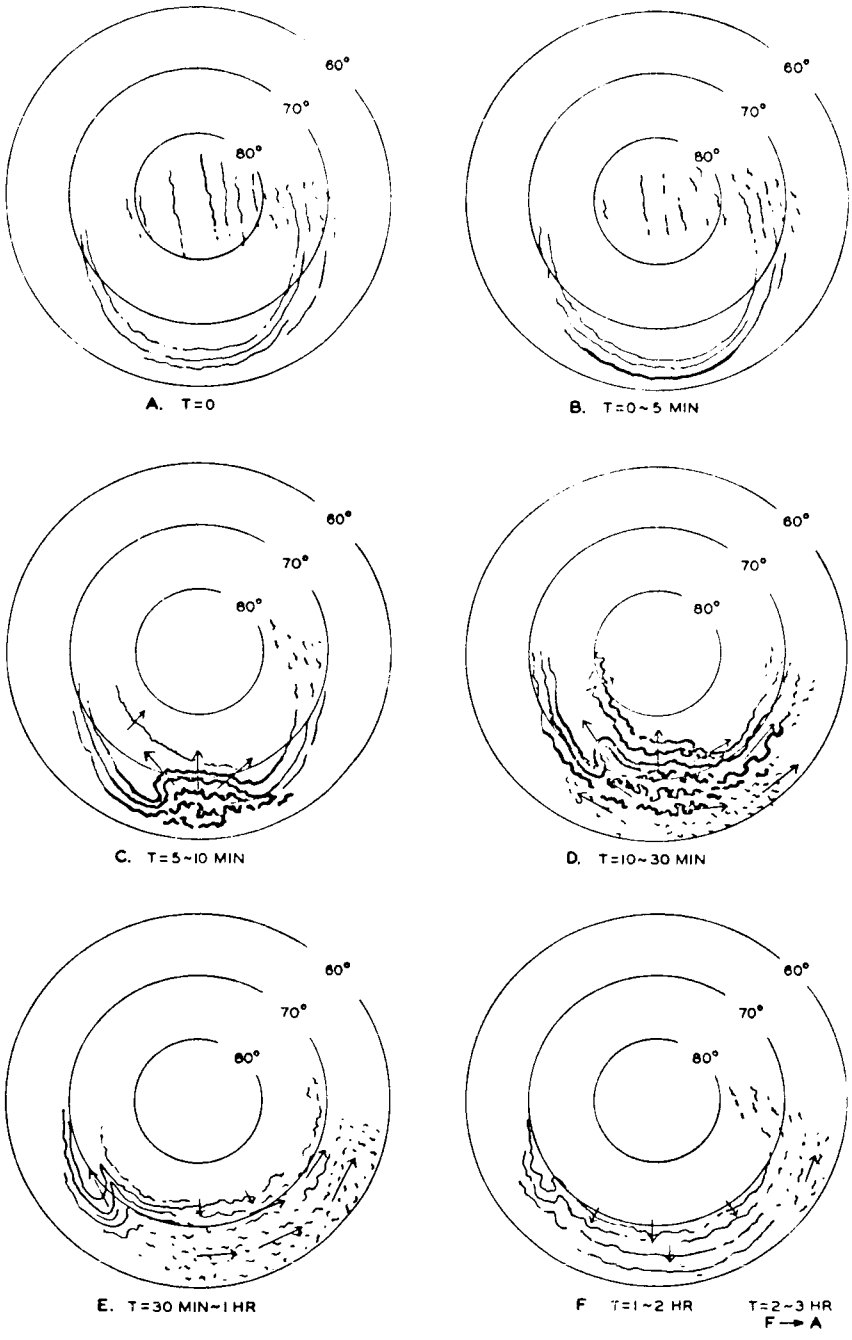


Figure 9–Spatial and temporal development of the auroral arcs (Akasofu, 1964).

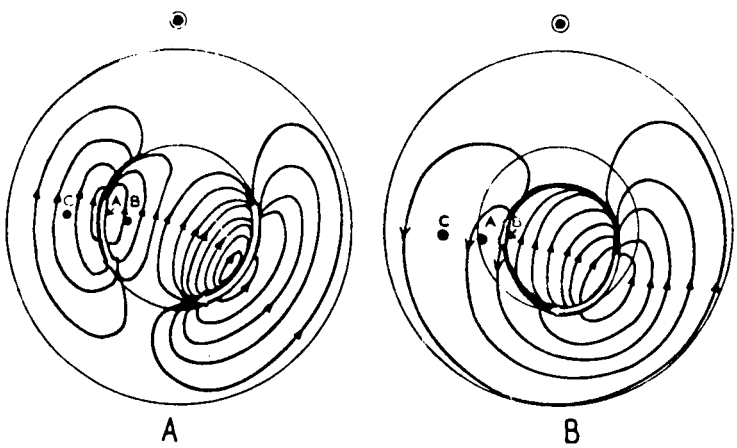


Figure 10–Schematic view of the equivalent current systems for polar substorm. (A) Conventional Model (B) New Model proposed by Akasofu (1965), in which the westward current flows along the auroral oval.

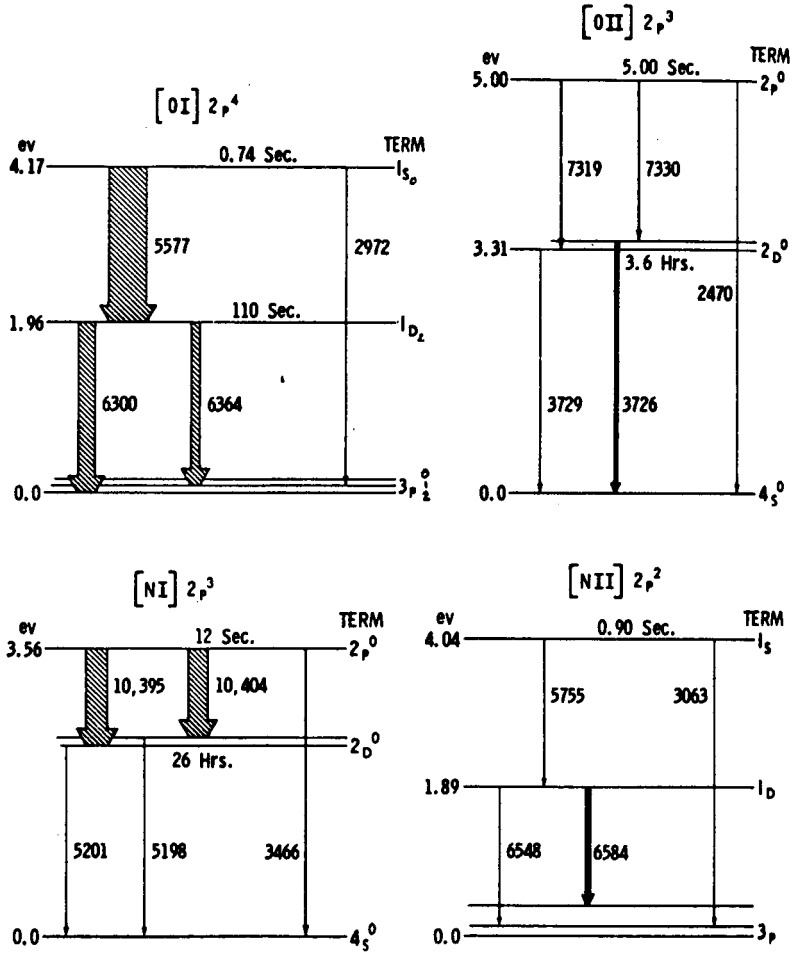


Figure 11–Energy level diagram of neutral and ionized atoms of oxygen OI, OII, and nitrogen NI, NII. Main transitions (all forbidden) are indicated by widths proportional to relative intensities. (Davis, 1965.)

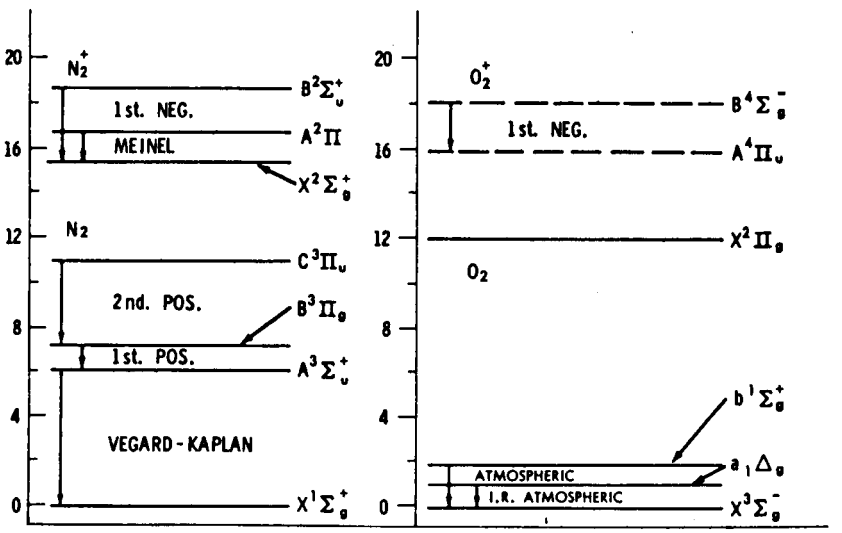


Figure 12–Main transition lines of molecular nitrogen and oxygen, and positive ions N_2^+ and O_2^+. (Davis, 1965.)

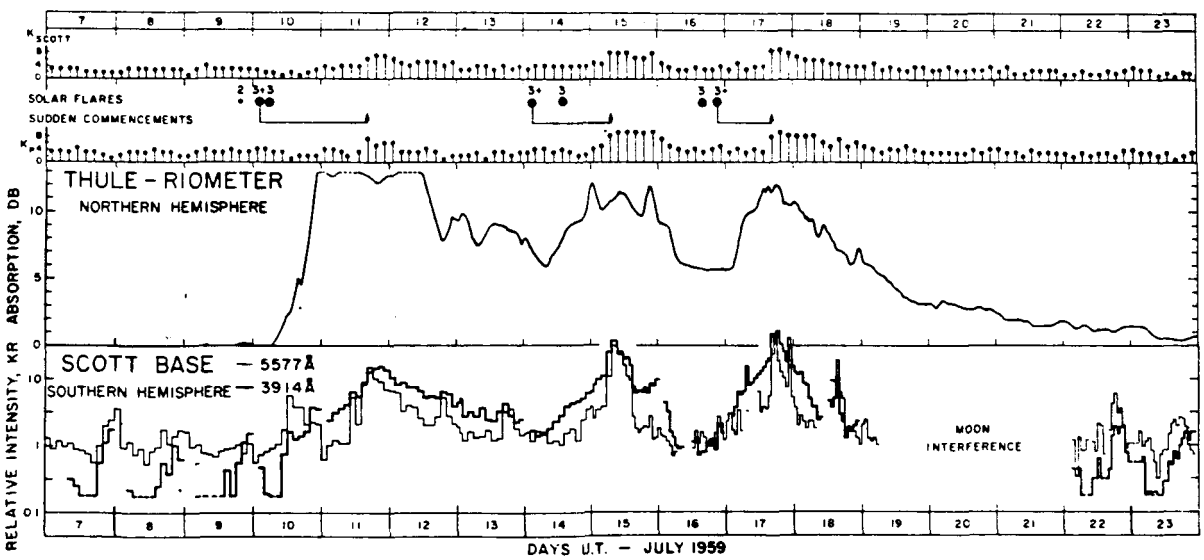


Figure 13–Time variation of polar cap glow aurora during 7-23, July 1959 reported by Sanford (1962), compared with K-index of Scott base in Antarctica, Solar flare and planetary K_p index, Thule-riometer records (λm 88°N) and intensities of zenith auroral emission (in KR); thin line 5566Å [OI], and thick line 3914Å [N_2^+] line, observed at Scott station in Antarctica (79° S). It should be noted that riometer data is from the northern hemisphere, aurora data from the southern. (Sanford, 1962).

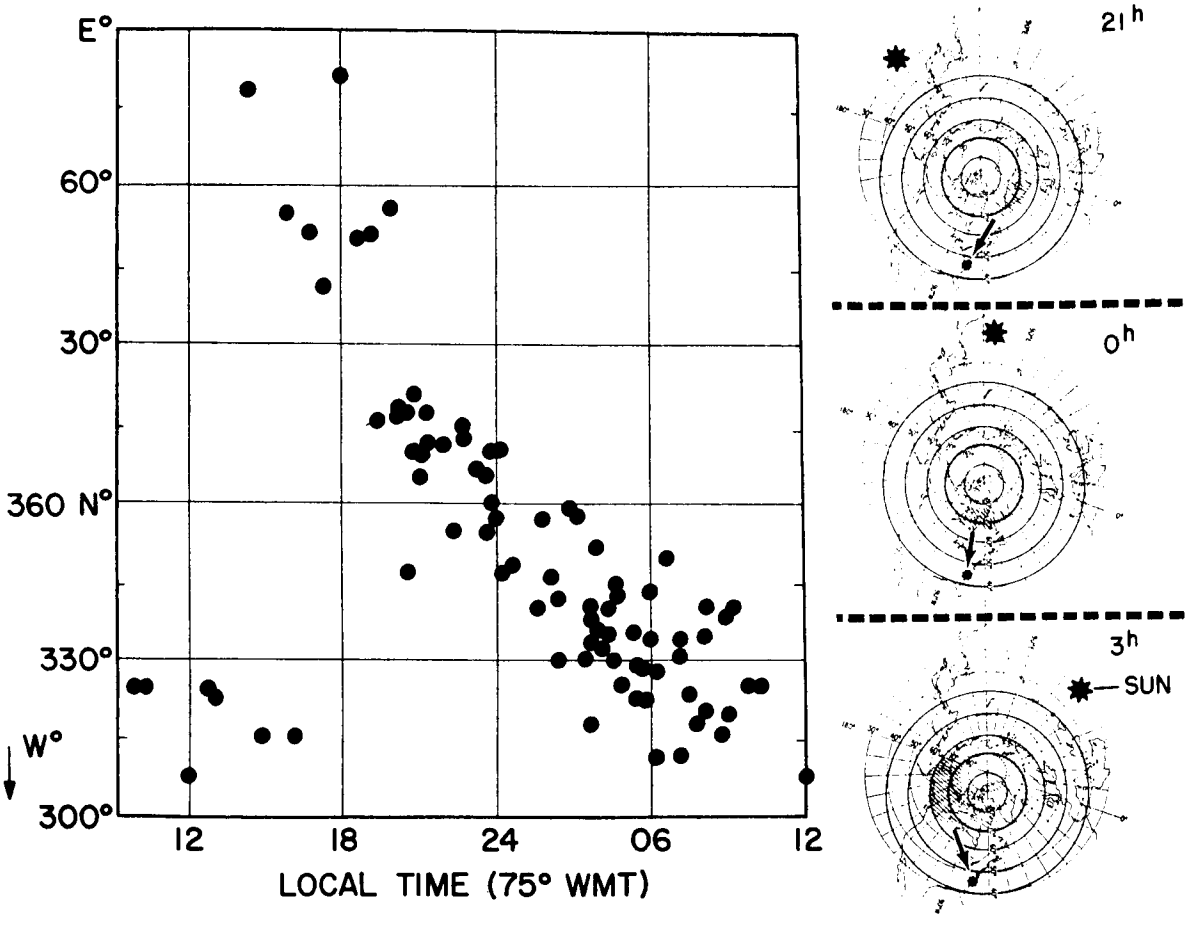


Figure 14–(a) Diurnal variation of arrival direction of infrasonic waves during magnetic storms observed at Washington, D. C. and (b) Shift of wave source (hatched area).

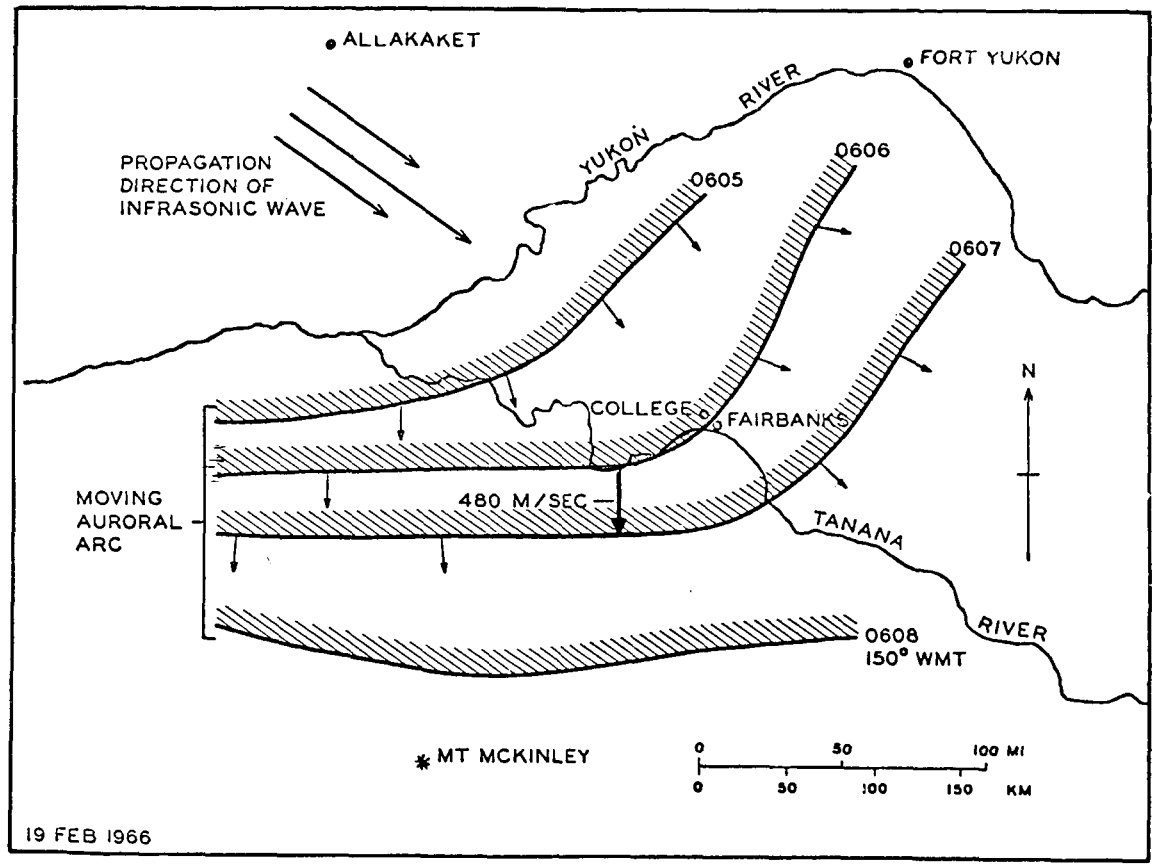


Figure 15–Horizontal projection of a fast-moving auroral arc at one-minute intervals, observed at College, Alaska, on February 19, 1966. (Time attached to arc is AST = UT – 10). (Wilson and Nichparenko, 1967).

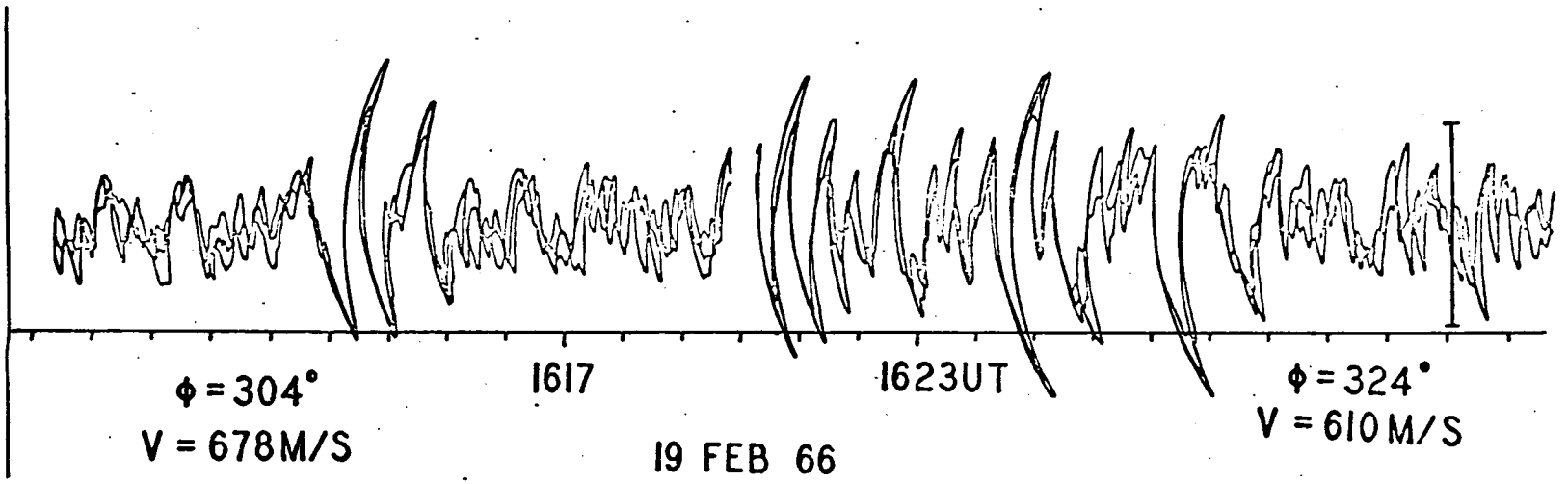


Figure 16–Superposed infrasonic records obtained by three stations in College, Alaska, on February 19, 1966. A vertical bar right indicates 5 dyne/cm^2 scale. (Wilson and Nichparenko, 1967).

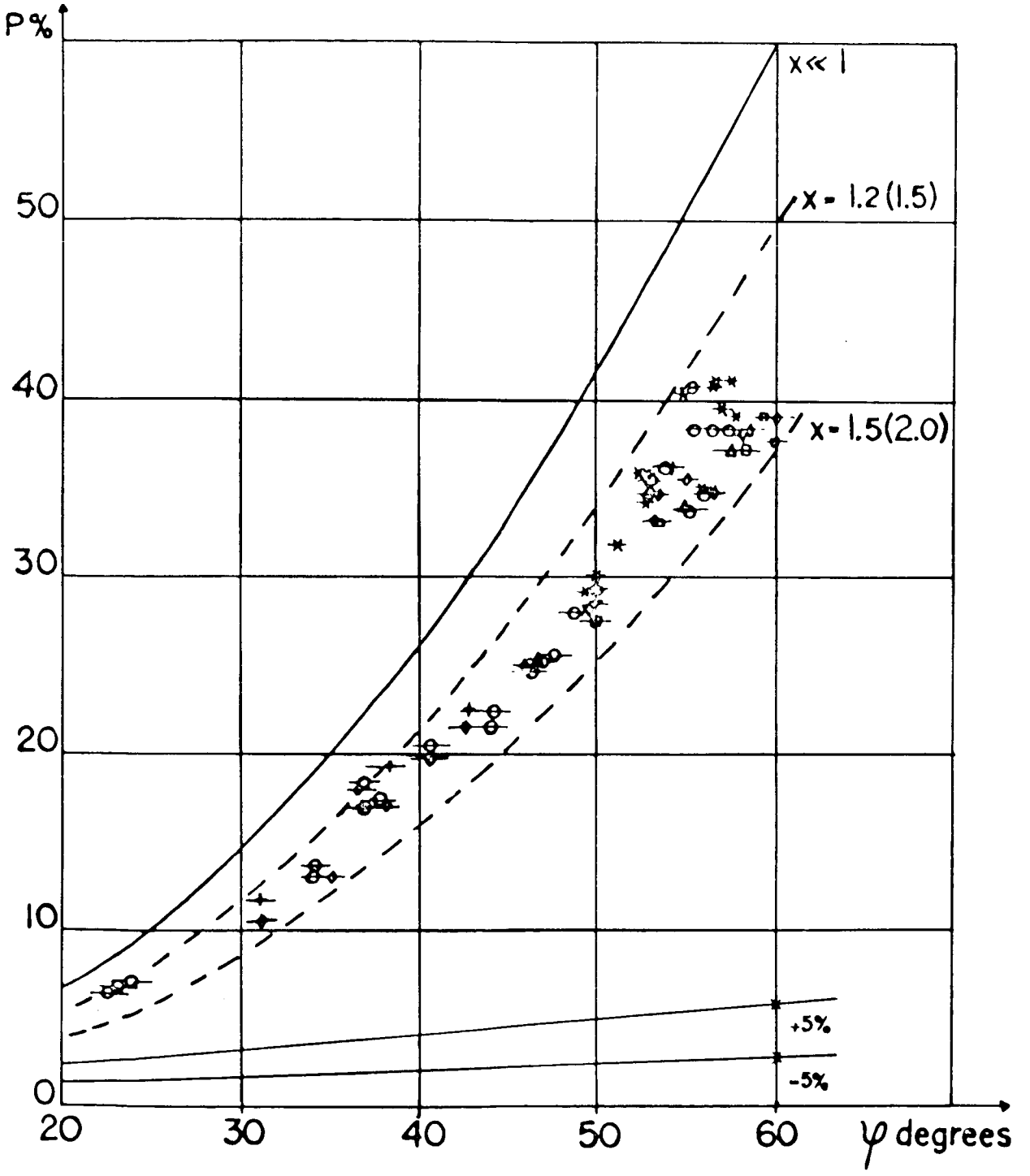


Figure 17–Observed and calculated polarization (percentage polarization, P) of blue light from noctilucent cloud versus scattering angle ψ. Full line corresponds to the Rayleigh scattering; dashed lines are calculated results based on Mie's theory. Four values of parameter x: x = 1.2 and 1.5 correspond to particle-radius r = 0.09 μ and r = 0.11 μ respectively, with n = 1.33 (water); x = (1.5) and (2.0) correspond to r = 0.11 μ and 0.16 μ with n = 1.55, respectively. n is the refractive index of particle. Two thin lines at the bottom indicate the magnitude of 5% error.(Witt, 1960.)

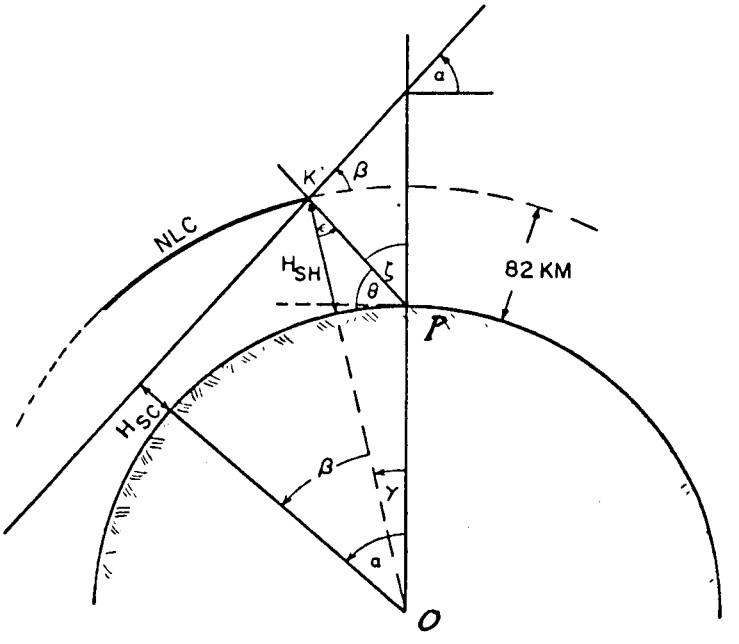
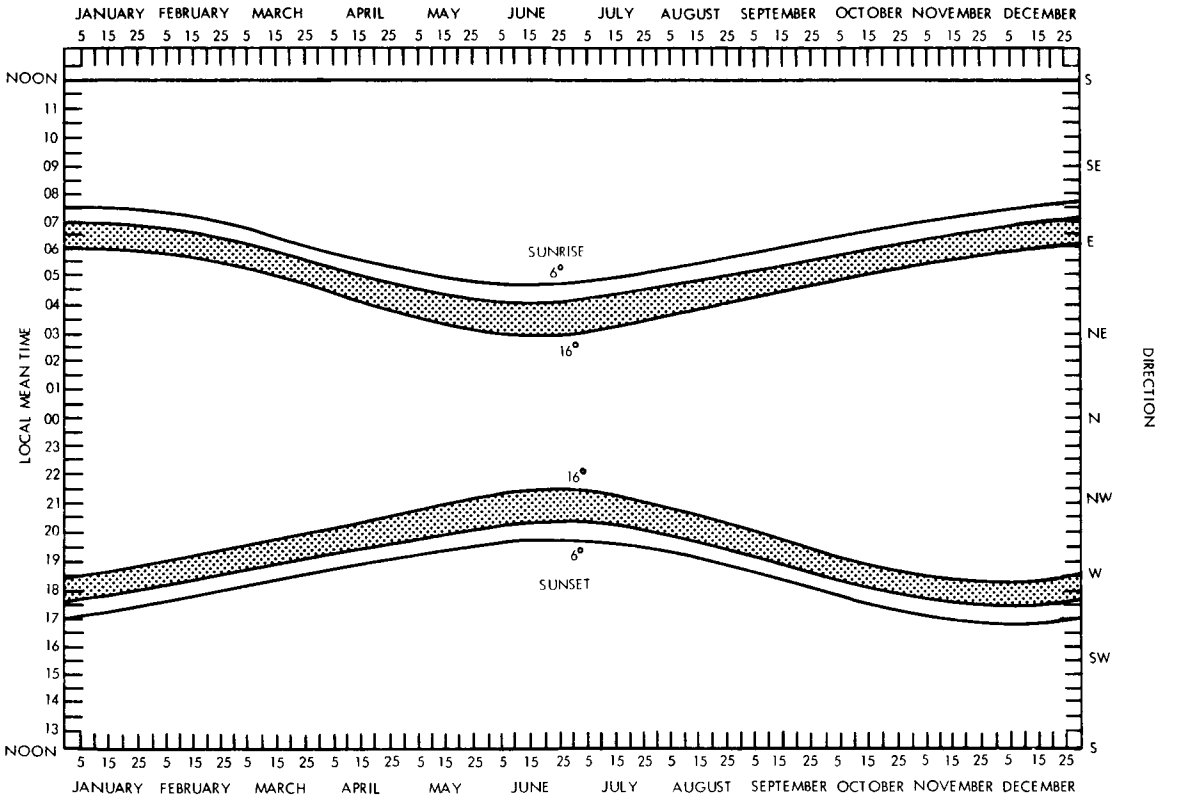


Figure 18–Geometric conditions illustrating conditions for seeing the noctilucent cloud (NLC) from the ground, where α is the solar depression angle, and H_{sc} is atmospheric screening height. (Fogle, 1966).



NOCTILUCENT CLOUD OBSERVATION PERIODS (SHADED AREA) FOR 40°N.

Figure 19(a)—The observable period of noctilucent cloud at different (geographic) latitudes, 40°.

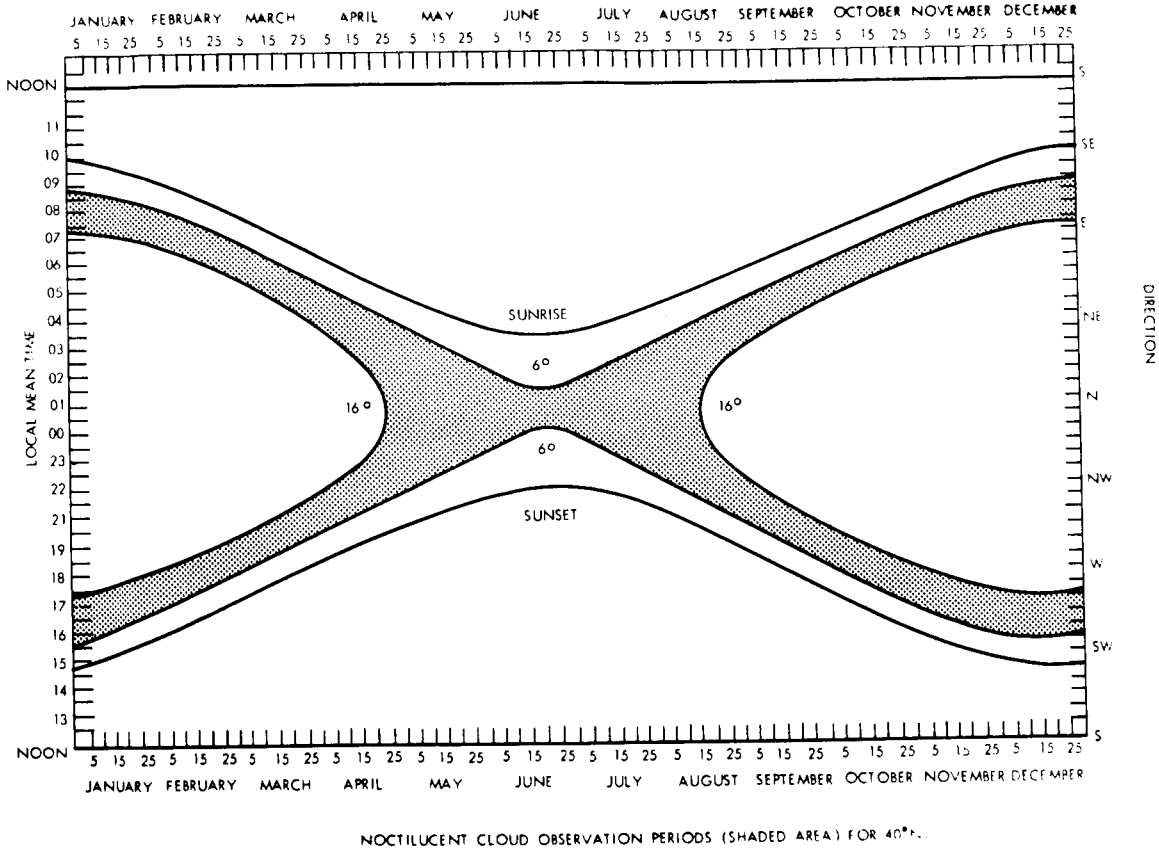
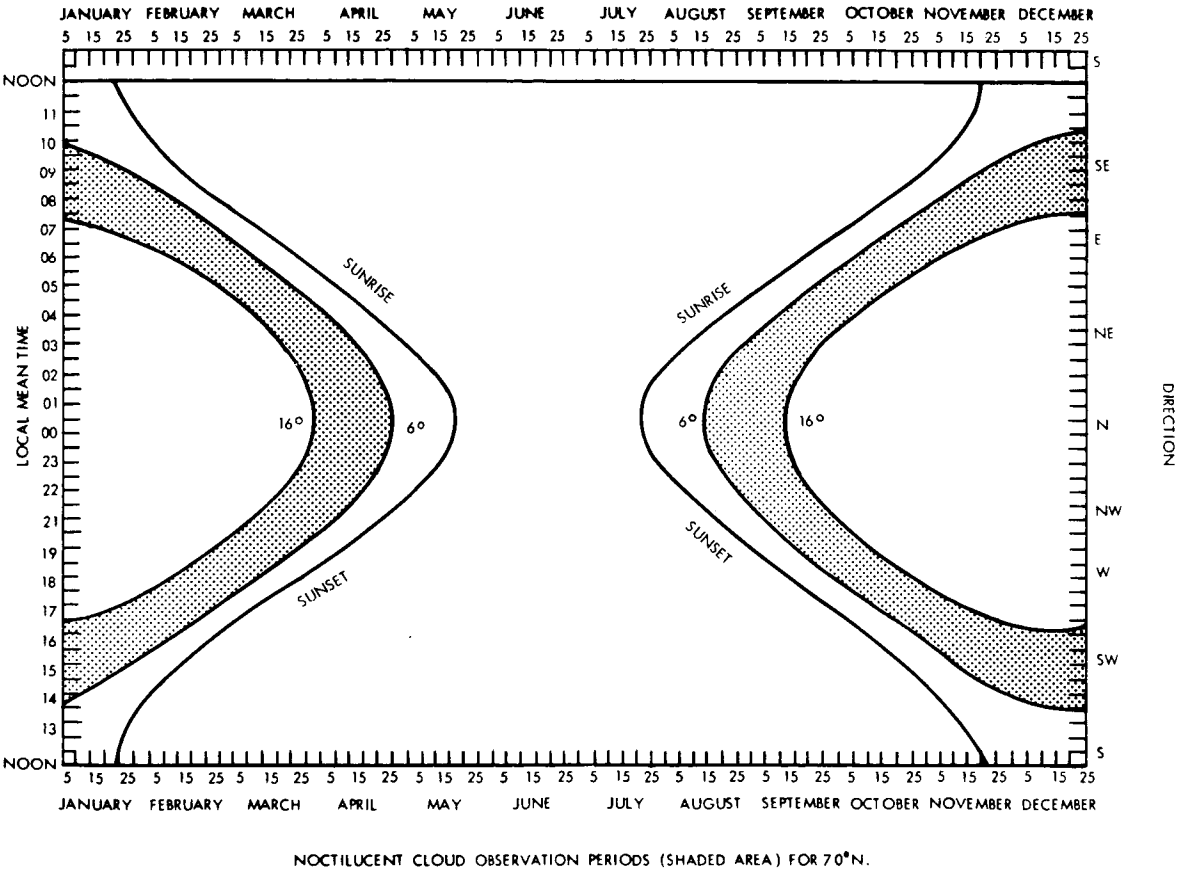


Figure 19(b)–The observable period of noctilucent cloud at different (geographic) latitudes, 60°.



NOCTILUCENT CLOUD OBSERVATION PERIODS (SHADED AREA) FOR 70°N.

Figure 19(c)—The observable period of noctilucent cloud at different (geographic) latitudes, 70°.

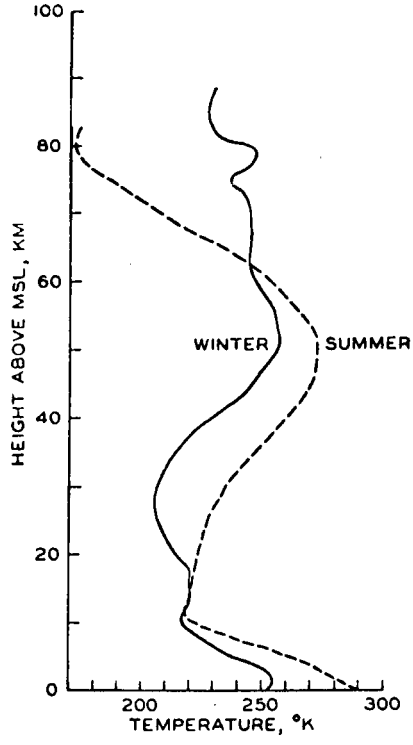


Figure 20–Vertical temperature distributions observed by rocket (grenade) soundings at Ft. Churchill ($\lambda = 59°N$, $\lambda_m = 68°.6N$) Canada. (Stroud et al., 1960).

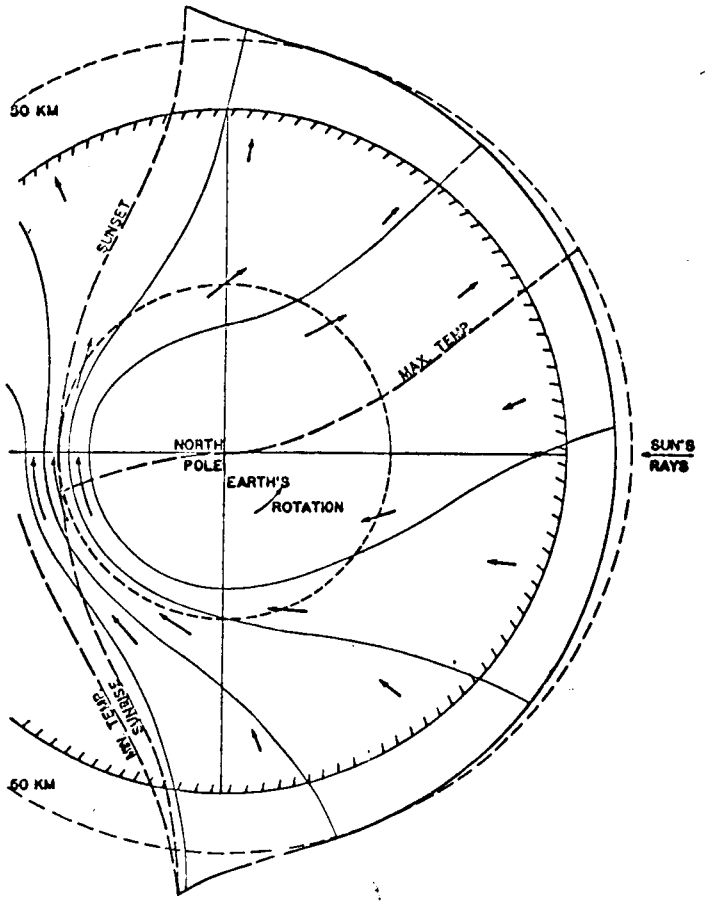


Figure 21—The stratospheric diurnal circulation at the summer solstice, projected on the equatorial plane. Thin solid curves are contours of constant temperature or thickness; the arrows indicate the tidal wind field. This diurnal circulation is imposed on the general summer easterly and weak poleward circulation (Webb, 1966 b).

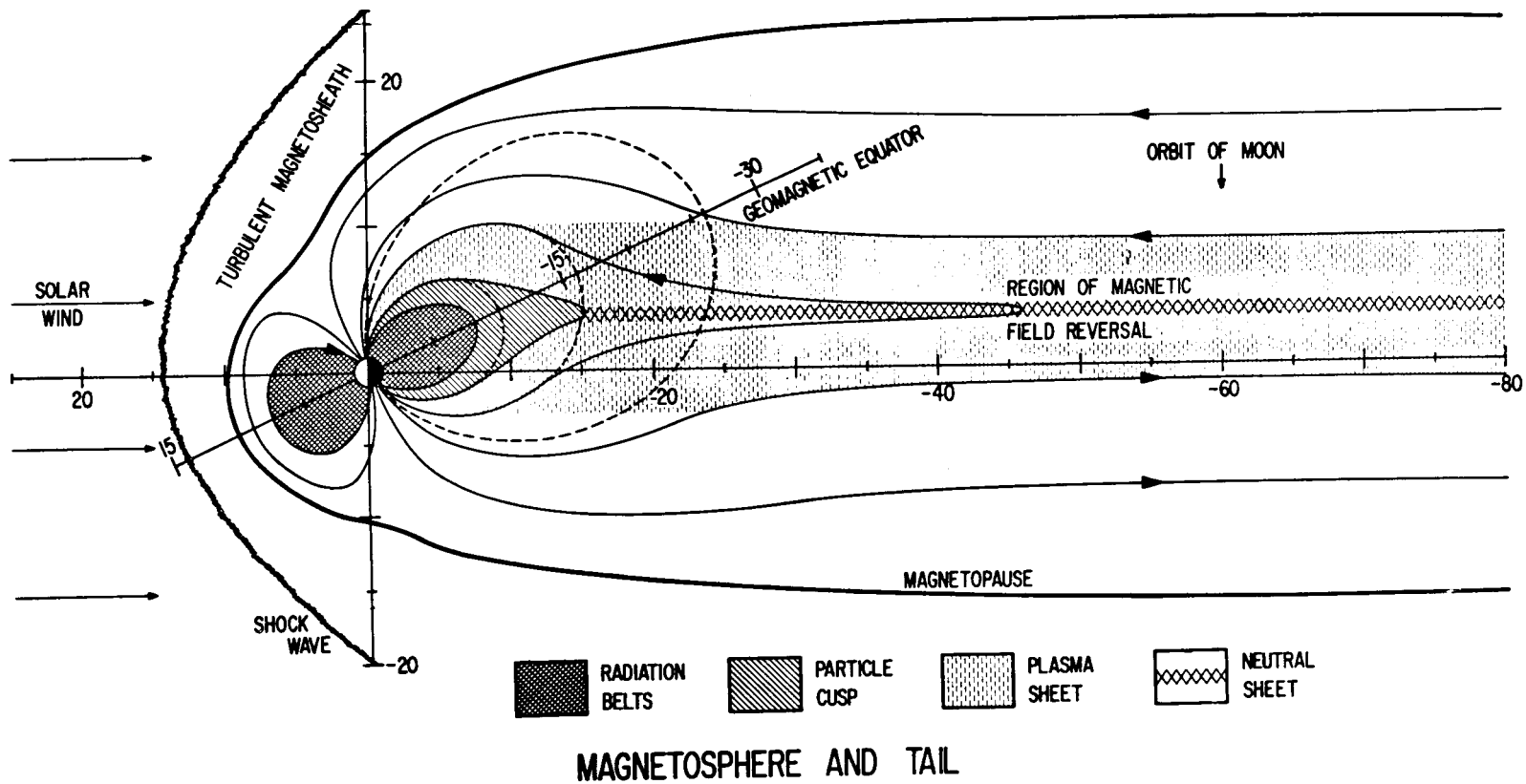


Figure 22–Cross section of the earth's magnetosphere and tail in noon-midnight meridian (Ness, 1966).