Terrestrial Magnetic Variations and Aurorae

E. O. HULBURT
Naval Research Laboratory, Washington D. C.

CONTENTS

INTRODUCTION

THE magnetic field of the earth is observed to undergo slow and rapid changes, the former measured in years and the latter in days, hours or minutes. The magnitudes of the magnetic variations are always small from a world-wide point of view for they rarely amount to more than a few percent of the total magnetic field of the earth. Gaussian harmonic analysis has shown that the slow variations are due to changes of some sort in the crust or the interior of the earth and that the more rapid changes arise from influences in part external to the surface of the earth. The influences are not in the low atmosphere and therefore, if on the earth at all, must be in the high atmosphere.

The rapid changes are of two types, periodic and erratic. The periodic types, associated with solar and lunar days, are known as the solar and lunar diurnal variations of the terrestrial magnetic field. The erratic types, known as magnetic disturbances or storms, are thought to be associated mainly with solar activities, as sunspots, faculi and eruptions; certain of the erratic types may possibly be due to extra-solar causes, as meteors and the sweep-up of material by the earth in its travel through space.

Although interest in these subjects was not dormant no startling advances were made in the early part of the present century. A few years ago, however, there occurred a recrudescence of activity which continues vigorously at the present time. This was due in large part to the recognition of electric and magnetic effects old in the literature of physics but new in the application to the high atmosphere; to the proposal, at first sight somewhat novel but now partially verified by experiment, that magnetic storms and aurorae were caused by bursts of ultraviolet light from the sun; and to the results of the recently instituted experimental investigations of the ionosphere by means of radio waves.

The following paragraphs present a summary of the more important facts of magnetic variations and the several theories which have been advanced in explanation. Since aurora and magnetic disturbance are in general closely connected a discussion is given of the facts and theories of the aurora. By the same token the allied subjects of the light of the night sky, the zodiacal light and the solar corona should also be treated; this, however, is not done for one must draw the line somewhere.

All theories of terrestrial magnetic variations assume in common that the variations arise from the effects of electrical charges and differ in their choice of effect and the situation of the charges. Three effects are recognized, dynamo action, diamagnetism and gravitational magnetic drift currents. The more important theories place the charged particles in the upper atmosphere. Other theories, for the most part inchoate or discarded, place the charged particles at distances of several earth's radii from the earth; such particles are not truly in the upper atmosphere.

Section A. Meteorology, Electricity and Magnetism of the Upper Atmosphere

Meteorology of the high atmosphere

The calculations of Maris[1] of the rates of diffusion of the atmospheric gases show that because of winds the atmosphere is uniformly mixed up to levels of about 150 km. Therefore, except for water-vapor and ozone, the gaseous composition of the atmosphere is the same up to about 150 km as it is at sea level. The drift and distortion of meteor trains[2] give direct evidence of winds and air currents up to 110 km. Above 150 km the distribution of the gases according to equilibrium under gravity may be expected, the proportion of the lighter gases increasing with height. The diffusion calculations were supported by the results of recent high altitude balloon flights.[2a] Samples of air at various levels enabled the determination of the amounts of oxygen in the atmosphere up to 29 km, of helium up to 21 km and of nitrogen up to 21.5 km. It was found that the proportions by volume of the respective gases were constant within a few percent from the ground to all heights reached. Deviations from constancy, small and variable, were observed. These remain to be explained, but appear to be attributable to irregularities in the lower atmosphere such as weather.

Observations[3] with sounding balloons to 33 km show that the temperature of the atmosphere increases from 219°K at 20 km to 226°K at 33 km. Above 33 km all estimates of the temperatures are theoretical. The calculations of Maris, Gowan[4] and others[5] as to the possible warming of the high atmosphere by the absorption of solar and earth radiation may be summarized thus: The calculations based on reasonable hypotheses indicate that the high atmosphere is warm, but are unable to say how warm, or to decide whether it remains warm or grows cool at night.

On the basis of a day temperature of 360°K above 80 km with complete mixing from sea level to about 130 km and isothermal equilibrium above 130 km, Maris drew up tables of the molecular densities of the gases of the atmosphere to great heights which are generally accepted as the most reasonable at the present time. A portion of the tables is given in Table I, where n is the total molecular density and z is the height above sea level, assuming that oxygen and nitrogen are in the molecular state. If these gases are in the atomic state the values of n above $z=200$ km increase by an order of magnitude or more.[6] Above $z=400$ km collisions are few and ordinary gas laws begin to break down; this is the outer fringe of the atmosphere.

The temperature 360°K of the upper atmosphere used by Maris seemed extreme at the time

[1] Maris, Terr. Mag. and Atmos. Elec. **33**, 233 (1928); **34**, 45 (1929).

[2] See summary by Hulburt, Publ. Ast. Soc. Pacific **44**, 178 (1928).

[2a] Prokofiew and Goltzman, Nature **133**, 918 (1934); Paneth and Glückauf, Nature **136**, 717 (1935); Lepape and Colange, Nature **137**, 459 (1936); Regener, Nature **138**, 544 (1936); Shepherd, National Geographic Society—U. S. Army Air Corps Stratosphere Flight of 1935, Technical Papers, p. 117 (1936).

[3] Frankenberger, Ann. d. Hydrographie u. Maritimen Meteorologie **59**, 20 (1931).

[4] Gowan, Proc. Roy. Soc. **A128**, 531 (1930).

[5] Martyn and Pulley, Proc. Roy. Soc. **A154**, 455 (1936).

[6] Hulburt, Terr. Mag. and Atmos. Elec. **40**, 193 (1935); Proc. Inst. Rad. Eng. **23**, 1492 (1935).

but there are further indications that even higher temperatures prevail. Calculation showed[6] that the absorption of sunlight by molecular oxygen in the great absorption band from 1850 to 1250A may increase the temperature by 50°K per hour in levels from about 180 to 300 km. The existence of the F_2 region of ionization above 300 km is fairly definite proof that the values of n of Table I for z above 250 km are too small. Radio evidence has been interpreted as meaning that the warm atmosphere above 150 km cools at night.[6]

Ionization of the upper atmosphere

Fairly complete measurements[7] of the ionosphere by the radio echo method of Breit and Tuve[8] are available for New York, Washington and Huancayo, Peru, during the years of sunspot minimum 1933 and 1934. The data have been summarized and analyzed;[6] in the present instance we need only to mention certain salient facts. The ionospheric observations show that for the sun overhead the equivalent electron density y for E, F_1 and F_2 regions is about 1.5×10^5, 3×10^5 and 10×10^5 at about 100, 200 and 300 km, respectively, but yield no complete information about the ionization below, between and above these levels except that the values at the levels are not exceeded. These facts are illustrated in the dotted curve of Fig. 1. On the assumption, as yet unchallenged, that the upper atmosphere is electrically neutral, or approximately so, y is also the density of positive ions.

TABLE I. *Theoretical molecular density n and free path γ in the high atmosphere.*

z	n	γ
60 km	9.9×10^{15}	0.03 cm
80	1.0×10^{15}	0.25
100	1.6×10^{14}	1.5
120	2.6×10^{13}	9.6
140	4.4×10^{12}	57
160	7.7×10^{11}	320
180	1.4×10^{11}	1800
200	2.5×10^{10}	10^4
250	2.5×10^{8}	10^6
300	7.7×10^{6}	10^7
350	7.3×10^{5}	10^8
400	4.1×10^{5}	10^9

A radio echo measurement of a region of ionization yields the lowest frequency f_c which just penetrates the region. The relation between f_c and the density of charged particles derived by equating to zero the refractive index of the medium for waves of frequency f_c, is

$$f_c^2 = e^2c^2y/\pi m, \tag{1}$$

where m and e e.m.u. are the mass and charge of the charged particle and c is the velocity of light. The charged particles are assumed to be singly charged, hence e is the electronic charge. If there are various types of ions the right-hand side of (1) is summed over all types. (1) is valid in the absence of a magnetic field; for the case of a magnetic field the formula is more elaborate and need not be given here. Therefore from (1) an experimental determination of f_c gives the ratio y/m, and leads to an evaluation of y only if m is known. For example, if f_c is observed to be 5000 kilocycles sec.$^{-1}$ we find from (1) that y is 3.1×10^5 for electrons and 1.24×10^{10} for ions of mass 3.6×10^{-23} grams or 40,000 times the mass of the electron.

Two scales of abscissas are marked in Fig. 1, the lower scale in terms of electron density and the upper scale in terms of ion density assuming that the mass of the average atmospheric ion is 3.6×10^{-23} grams. The observed[7] magnetic splitting of the radio echoes returned from the ionosphere indicates that above 200 km, i.e., for F_1 and F_2, the ionization is largely electronic. Hence in Fig. 1 for the part of the dotted curve above 200 km the lower scale applies. Experiment has not yet determined definitely whether E region is mainly electronic or ionic; at present a majority of experiments indicate ions and a few experiments electrons. Therefore, for the part of the curve below 200 km either of the abscissa scales may be used, or a scale representing a suitable mixture of ions and electrons. As brought out later the matter is of vital importance to the theories of magnetic variations, for if future experiment proves that the ionization below 200

[7] The comprehensive ionosphere observations of the National Bureau of Standards, of the Department of Terrestrial Magnetism, Carnegie Institution, and of the Bell Telephone Laboratories, have appeared in a number of papers some of which are: Kirby, Berkner and Stuart, Proc. Inst. Rad. Eng. **22**, 481 (1934); Berkner and Wells, Proc. Inst. Rad. Eng. **22**, 1102 (1934), Terr. Mag. and Atmos. Elec. **39**, 209 (1934); Gilliland, Nat. Bur. Standards J. Research **14**, 283 (1935); Kirby and Judson, Nat. Bur. Standards J. Research **14**, 469 (1935); Schafer and Goodall, Proc. Inst. Radio Eng. **19**, 1434 (1931); **20**, 1131 (1932); **23**, 1670 (1935).

[8] Breit and Tuve, Phys. Rev. **28**, 554 (1926).

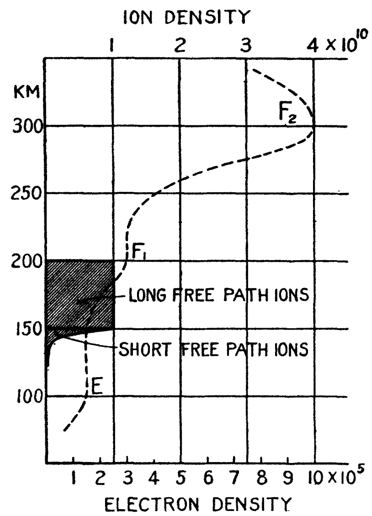

FIG. 1. Dotted curve, ionosphere at equatorial noon 1933–1934, sketched from observation;[7] shaded area, ion banks assumed in terrestrial magnetic theory.[9]

km is predominately electronic all of the theories will fall to the ground or require fundamental modification.

It was shown[6] that the observed values of y, the maximum point on the y,z curve, for E during 1933 and 1934 conformed to the relation, derived from the hypothesis that the ionization is caused by the ultraviolet light of the sun,

$$y = y_0(\cos \zeta)^{\frac{1}{2}}, \tag{2}$$

where ζ is the zenith angle of the sun and y_0 is the value of y for the sun overhead, i.e., $\zeta = 0$. (2) is valid for ζ from 0° to about 85°. For F_1 and F_2 the observed values of y departed considerably from (2), but a reasonable theory was proposed[6] which accounted for the departures. Thus there is at present no objection to the hypothesis that the ionization over the day hemisphere is expressed approximately by (2). The average ionization during the night is about 1/10 of the noon value.

Comparing the observed value 10^6 of the electron density in F_2 region with the molecular densities above 300 km of Table I it is seen that they are of the same order of magnitude. Therefore the outer fringe of the atmosphere is practically completely ionized. The conclusion is arrived at from theory easily enough, for the sun is such a powerful ionizing agent, but the realization that it has been demonstrated by direct experiment leaves one a little breathless.

Long and short free path regions

Two regions of the high atmosphere are differentiated,[9] the short free path region and the long free path region. In the first region the radius r of magnetic gyration of the ions or electrons is large compared to the mean free path γ; in the second r is small with respect to γ. r is given by

$$r = mu/He, \tag{3}$$

where u, m and e are the velocity of temperature agitation, mass and charge, respectively, of the charged particle and H is the component of the earth's magnetic field perpendicular to u. The free path γ is, respectively, for ions and electrons

$$\gamma = 1/\sqrt{2}\pi n a^2 \quad \text{and} \quad \gamma = 4/\pi n a^2, \tag{4}$$

where n is the molecular density and $a = 3 \times 10^{-8}$ cm is the kinetic theory diameter of the average atmospheric molecule or atom.

γ for ions from (4) is given in the last column of Table I. From (3) r is 20 and 1.3 cm for ions and electrons, respectively, for $H = 0.3$ gauss as at the equator. The ratio γ/r in the case of ions becomes large above 150 km, for electrons it is large above 100 km. Therefore at the equator the short free path region merges into the long free path region at about 150 km for ions and 100 km for electrons. In temperate latitudes $H = 0.5$ gauss, and, for a temperature of 360°K, the respective levels are about 10 km lower. If oxygen and nitrogen of the high atmosphere are in the atomic state n is greater than the values of Table I and the foregoing levels are 10 to 30 km higher.

Conductivity

The conductivity σ per unit volume of an ionized gas is in e.m.u.

$$\sigma = \sum y e^2 \gamma / 2mu, \tag{5}$$

where y is the density of charged particles; the summation is taken over all classes of charged particles. In a magnetized ionized gas the conductivity along H is undisturbed by the magnetic field and is given by (5). The conductivity σ' across the field is approximately[10]

---

[9] Hulburt, Phys. Rev. **31**, 1018 (1928); **34**, 1167 (1929); **35**, 240 (1930); **39**, 977 (1932).

[10] Pederson, *The Propagation of Radio Waves* (1927), Chap. VII.

$$\sigma' = \sigma/(1+\gamma^2/r^2). \qquad (6)$$

A more rigorous calculation by Page[11] gave

$$\sigma' = \sigma(1+2\gamma^2/r^2)/(1+\gamma^2/r^2)^2. \qquad (7)$$

Eq. (6) was derived by averaging the effect of collisions in shortening the free magnetic spiraling of the ions and (7) by the same averaging process and taking into account motion of the ions in the direction of the impressed field. They differ but slightly and both yield $\sigma'=0$ and σ in the long and short free path regions, respectively.

Motion of ions in crossed fields, gravitational magnetic drift

Let an ion of charge e and mass m in a magnetic field H be acted on by a constant force X. X and H are along the positive directions of the x and z axes, respectively. In c.g.s. e.m.u. the equations of motion of the ion are

$$\begin{aligned} m\ddot{x} &= X + He\dot{y}, \\ m\ddot{y} &= -He\dot{x}, \\ m\ddot{z} &= 0. \end{aligned} \qquad (8)$$

Let $\omega = -He/m$ and for $t=0$, $\dot{x}=\dot{x}_0$, $\dot{y}=\dot{y}_0$, $\dot{z}=\dot{z}_0$. The solution of (8) yields

$$\begin{aligned} \dot{x} &= \dot{x}_0 \cos \omega t + (\dot{y}_0 + X/He) \sin \omega t, \\ \dot{y} &= -X/He + (\dot{y}_0 + X/He) \cos \omega t - \dot{x}_0 \sin \omega t, \\ \dot{z} &= \dot{z}_0. \end{aligned} \qquad (9)$$

From (9) the ions move in cycloidal paths with a drift in a direction perpendicular to X and H. If the mean values of $\dot{x}_0$, $\dot{y}_0$, $\dot{z}_0$ are zero the average of $\dot{x}$, $\dot{y}$, $\dot{z}$ with respect to time is, since the periodic terms vanish,

$$0, \; -X/He, \; 0. \qquad (10)$$

Therefore, the ions drift[12–14] along Y with a constant velocity X/He. The averaging eliminates correctly the kinks in the cycloidal paths of the ions.

At the equator, where H is horizontal and north, the ions are pulled downward by gravity and from (10) drift in a horizontal direction with a velocity v given by

$$v = mg/He. \qquad (11)$$

[11] Page, Phys. Rev. **34**, 763 (1929).
[12] Maris and Hulburt, Phys. Rev. **33**, 412 (1929).
[13] Chapman, Proc. Roy. Soc. **A122**, 369 (1929).
[14] Page, Phys. Rev. **33**, 553, 823 (1929).

The positive ions move eastward and the negative particles westward, constituting an eastward electric current termed "gravitational magnetic drift current." At 200 km altitude at the equator $g=920$ cm sec.$^{-2}$, $H=0.29$ gauss and v is 11.8, 5.3 and 1.3 cm sec.$^{-1}$ for nitrogen molecules and oxygen and helium atoms, respectively.

From (10) an electric field E normal to H causes ions of both signs to move perpendicularly to E and H with the same velocity

$$v = -E/H. \qquad (12)$$

For E eastward ions at the equator drift upward, and downward for a westward E. The drift is not an electric current since ions of both signs move in the same direction with the same velocity.

The gravitational magnetic drift (11) and the electric magnetic drift (12) are fully effective only in the long free path region and are zero in the short free path region. It may be shown after the manner of Pederson[10] that they vary with the free path approximately according to the ratio

$$1/(1+r^2/\gamma^2). \qquad (13)$$

Dynamo effect

The dynamo effect is the electromotive force E induced in a conductor moving with velocity v across a magnetic field H. In c.g.s e m.u.

$$E = -vH, \qquad (14)$$

where E, v and H are the components along the positive directions of the x, y and z axes, respectively. This follows from Faraday's law of electromagnetic induction. The short free path ionized region of the upper atmosphere is a conductor and if, due to winds, thermal expansion or tidal forces, it moves across H electromotive forces according to (14) are developed. If conducting paths exist the electromotive forces give rise to currents, and these in turn to magnetic effects.

Motions of the long free path regions across H also develop E in accord with (14). For the force which sets the region in motion causes, by (10), a drift of positive and negative charges equivalent to the movement of charges required to yield E of (14).

Diamagnetism

The long free path region of the upper atmosphere is diamagnetic.[15] The intensity of magnetization i of free charges, as ions and electrons, of density y is

$$i = -ykt/H, \tag{15}$$

where t is the temperature Kelvin and $k = 1.372 \times 10^{-16}$ erg deg^{-1}. Just as in the case of the drift currents i varies with the free path according to (13) being zero in the short free path region.

Section B. Solar Diurnal Variation of Terrestrial Magnetism

Observed facts of the solar diurnal variation

The average daily variation, denoted by S, of the terrestrial magnetic field for various geographic latitudes is given in Fig. 2, in which N, V and W, the components of S in the north horizontal, vertical and west horizontal directions, respectively, are plotted against local time.[16] For latitudes below 40° N increases to a maximum during the daylight hours and for latitudes above 40° to a minimum. The day maximum or minimum of N occurs from 1 to 2 hours before noon or between 10 and 11 A.M. The amplitude of S increases with sunspots being nearly twice as great at sunspot maximum as at sunspot minimum; N is about 30γ (γ is 10^{-5} gauss) and 15γ, respectively, for the two epoques. Harmonic analysis indicates that about three-fourths of S arise from overhead causes and one-fourth from induced currents in the earth.[17]

Dynamo theory

The dynamo theory of S, originally proposed by Balfour Stewart, was extended in considerable detail by Schuster.[17] The theory conceived that S was caused by horizontal movements of conducting regions of the high atmosphere across the vertical component of the terrestrial magnetic field. The motion induced electromotive forces according to (14) which gave rise to electrical currents and hence to the magnetic variations of Fig. 2. The type of atmospheric

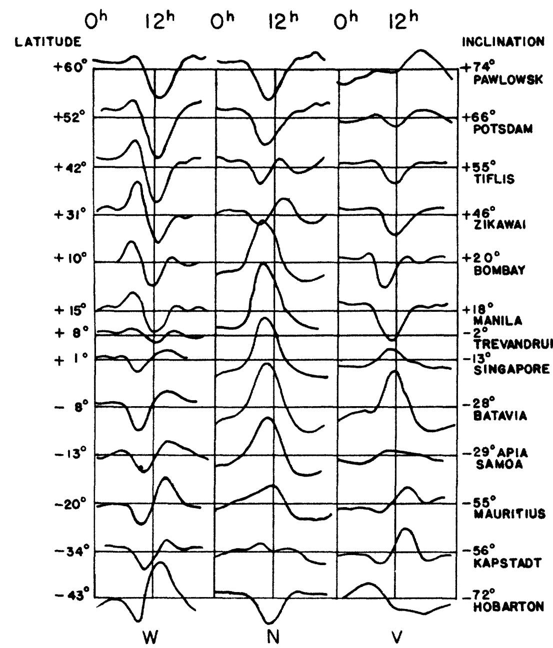

FIG. 2. Daily variation of terrestrial magnetic elements at equinox.[16] The inclination gives the magnetic latitude.

current necessary to cause S, worked out by Bartels,[18] is shown in Fig. 3; the numbers along the lines of current flow are in 10^3 amperes. Thus the figure indicates in the northern hemisphere in the day a current sheet of 62,000 amperes wheeling counterclockwise around a center at about 40° latitude and 11 A.M. and at night a current sheet of 32,000 amperes wheeling in the opposite direction, or an excess of day over night of 94,000 amperes. An epoque midway between maximum and minimum sunspots is referred to. The maximum current density occurs at 11 A.M. at the equator and has the value 3×10^{-5} c.g.s. e.m.u. excess day over night value.

The horizontal movements of the high atmosphere were calculated from the observed daily fluctuation of the barometric pressure. Two components of barometric variation are recognized, a regular component of 12 hour period and of amplitude 1 mm of mercury at the equator and a less regular, erratic component of 24 hour period and amplitude about 0.3 mm. The origin of the two components, whether thermal or tidal, has been much discussed.[19] If

[15] Gunn, Phys. Rev. **32**, 133 (1928); Terr. Mag. and Atmos. Elec. **34**, 17 (1929).
[16] Angenheister, *Handbuch der Physik* **15**, 302 (1927).
[17] Schuster, Phil. Trans. Roy. Soc. **A180**, 467 (1889); **208**, 163 (1908).
[18] Bartels, *Handbuch Exp. Physik* **25**, 640 (1928).
[19] Humphreys, *Physics of the Air* (1929), Chap. 14.

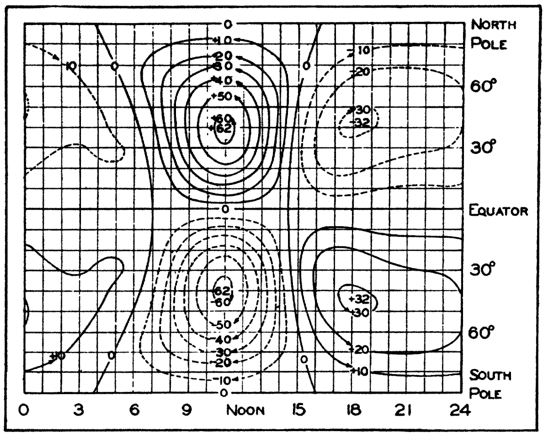

FIG. 3. Theoretical current system in the upper atmosphere necessary to produce solar diurnal variation in the magnetic field of the earth;[18] from magnetic data of 1902, an epoque of sunspot minimum.

the solar semidiurnal component is tidal some new hypothesis must be brought in, as resonance of the free atmosphere,[20] to account for the fact that it is 15 times larger than the lunar semidiurnal barometric variation, of amplitude 0.063 mm, whereas the tide raising force of the sun is only 2/5 that of the moon. The solar diurnal component may reasonably be supposed to be a thermal effect and to result from a redistribution of the atmosphere caused by winds which blow from expanding regions heated by the sun.

Qualitatively, winds blowing away in all directions from a high pressure area at noon at the equator will induce current sheets of the type of Fig. 3. For in day north latitudes above 40° northward velocities across the vertical component of H induce a westward e.m.f. which drives the circular current sheet of Fig. 3, provided the reasonable assumption is made that the conductivity is a maximum under the sun and decreases with increasing zenith angle of the sun to small values at night. Similarly a 12 hour tidal movement of the atmosphere will induce current systems of the type of Fig. 3 with components of 12 and 24 hours, the relative amplitudes of the components depending on the relation assumed for the dependence of the high atmospheric conductivity on the zenith angle of the sun.

[20] Chapman, Quarterly J. Roy. Met. Soc. **45**, 113 (1919); Nature **103**, 185 (1919).

As the result of an extensive mathematical analysis Schuster found that the electric currents inferred from the barometer readings, although of the correct type, disagreed in phase with those of Fig. 3. In fact they were nearly in opposition, yielding a maximum value of N at 2 to 4 P.M. in contrast to the observed time 10 to 11 A.M. A discrepancy was brought out between the observed and theoretical ratios of the amplitudes of the 24 hour and 12 hour periodic terms derived from harmonic analysis, the former being about 9 and the latter 3. The conclusion, not altogether unexpected, was drawn that if the dynamo hypothesis were to be retained at all the wind systems of the upper reaches of the atmosphere could not be inferred in a simple manner from sea level barometer readings. Discussion of more recent magnetic and meteorological data by Walker,[21] Chapman[22] and others confirmed Schuster's conclusion. The theory was left dangling.

The conductivity σ of the high atmosphere was calculated on the uncertain assumption that the 12 hour barometric variation was due to atmospheric tides of a simple nature such that the tidal motion is uniformly distributed through a vertical section of the atmosphere. The 1 mm amplitude of the variation may be taken to mean that $\frac{1}{2}\times 760$ of the atmosphere remains fixed with reference to the sun, and hence that there is a horizontal ebb and flow of the atmosphere with an average velocity of 1/1520 of the peripheral velocity of rotation of the earth. At the equator this amounts to an atmospheric tidal current of 30 cm sec.$^{-1}$ or 1 km hr.$^{-1}$.

From (14) we write $\sigma = i/vH$ and with $i=3\times 10^{-5}$, $H=0.3$ gauss and $v=30$, find

$$\sigma = 3\times 10^{-6} \text{ e.m.u.}$$

for equatorial noon, which is the value obtained by Schuster. However, it must not be supposed that the foregoing brief calculation represents adequately Schuster's more complete treatment; it merely presents correctly the physical ideas and yields a numerical value of the correct order.

If atmospheric resonance is assumed the air motions will differ from those which were calcu-

[21] Walker, Proc. Roy. Soc. **A89**, 379 (1914).
[22] Chapman, Phil. Trans. Roy. Soc. **A218**, 1 (1919); **225**, 49 (1926), and references *infra*.

lated from purely tidal effects in a nonresonating atmosphere and the above value of σ is meaningless. If the solar barometric variation is attributed to heating of the atmosphere by the sun, the atmospheric winds and the electrical conductivity required by the dynamo theory cannot be calculated unless the heating at all heights in the atmosphere is known or assumed.

The dynamo theory may be examined anew in the light of ionosphere measurements. Of the three daytime ionospheric regions only E contributes to the conductivity across H for F_1 and F_2 are in the long free path region. Referring to Fig. 1 assume that E ionization is constant from 100 to 150 km and is composed of equal numbers of electrons and positive ions of density y given by (2) with $y_0 = 1.5\times10^5$, which is the maximum number of electrons permitted by radio ionospheric measurements interpreted according to (1). The conductivity of E region may then be calculated over the day hemisphere. The electrons cannot move across H and the conductivity of a 1 cm^2 vertical column from 100 to 150 km is from (2), (5), (6) and Table I

$$\sigma = 7.2\times10^{-10} (\cos\zeta)^{\frac{1}{2}}. \qquad (16)$$

The average σ for day latitudes from 0° to 40° is about 5×10^{-10}; this is approximately four orders of magnitude below Schuster's value.

The total eastward current in Fig. 3 flowing between 0° and 40° latitude is 94,000 amperes excess day over night, or an average current density across a 1 cm^2 vertical column of E region of 2×10^{-5} e.m.u. To produce the current by dynamo action requires north moving winds in day latitudes above 40°, where the average vertical component of H is about 0.5 gauss, of velocity, from (14), $2\times10^{-5}/5\times10^{-10}\times0.5 = 8\times10^4$ cm sec.$^{-1}$ or 2900 km hr.$^{-1}$. The existence of such high winds appears improbable and we conclude that the dynamo theory fails under the assumptions which have been made.

An acceptable dynamo theory would require winds of reasonable velocity, say, less than 100 km hr.$^{-1}$. Therefore E conductivity must be at least 30 times greater than 5×10^{-10}. But E conductivity cannot be increased by assuming more electrons, for the maximum number of electrons permitted by radio ionospheric measurements has already been assumed. A way out is to appeal to the fact that the radio measurements as yet do not exclude the possibility of ions up to the density given by (2) with $y_0 = 6\times10^9$. If this density exists from 100 to 150 km σ is 4×10^{-5} and the breeze necessary to produce the 94,000 amperes is only 35 meters hr.$^{-1}$, which is a fairly gentle breeze. Thus there is some latitude between the limits at present set by E ionization observations, i e., electrons or ions, and the type of ionization which yields an E conductivity requisite to a reasonable dynamo theory. This was clearly recognized in early considerations of the ionosphere[9] in which an arbitrary selection of E electron and ion densities between the permissible limits was made.

The foregoing conclusion also applies to those portions of all theories of magnetic variation described later which attribute the magnetic effects to electric currents in conducting regions of the high atmosphere; reasonable causes of such currents cannot be proposed as long as the short free path ionization is predominately electronic.

Further pursuit of the dynamo theory entails a scrutiny of secondary effects to see either that they provide desirable explanations of observed facts or that they do not lead to conflict with observation. For example, north moving air in day latitudes above 40° induces a westward e.m.f. which together with H, by (12), causes the long free path ionization, F_1 and F_2, over the day hemisphere to drift with a component downward, and over the night hemisphere upward. These drift velocities are of the same magnitude as the north moving breeze. Therefore a dynamo theory which assumes, for example, a 40 km hr.$^{-1}$ breeze must be prepared to accept the consequences of a 20 to 40 km hr.$^{-1}$ downward drift of day F_1 and F_2 ionization and a similar upward drift of night F ionization.

We take leave of the dynamo theory with the remark that the theory has neither been proved nor disproved. To this extent it is a possible theory, but to be reasonably possible E ionization must be largely ionic.

Diamagnetic theory

Gunn[15] remarked that the ionization in the long free path region was diamagnetic according to (15) and worked out a diamagnetic theory of

S. Over the daylight hemisphere the intensity of magnetization i of a 1 cm^2 vertical column of the long free path region containing y' charged particles was assumed to be given by

$$i = i_0 \cos \zeta, \tag{17}$$

and at night i was assumed to be less than 1/10 of the noon value. The diamagnetism of such a distribution was shown[15] to produce a magnetic field in close accord with the observed curves of Fig. 2. A repetition of the calculations by Chapman[23] led to the same conclusion.

Introducing (15) into assumption (17) gives

$$y't = y_0't_0 \cos \zeta, \tag{18}$$

where y_0' and t_0 refer to $\zeta = 0$.

Ionosphere observations[7] secured after the development of the theory indicate[6] that $y' \sim (\cos \zeta)^{\frac{1}{2}}$ and support the view that t increases to a maximum near midday, and therefore are in good qualitative agreement with assumption (17).

Quantitative agreement with the observed magnitude of S, i.e., a horizontal component of H of external origin of about 15γ at equinoxial noon at the equator for an epoque midway between minimum and maximum of sunspots, required that y_0' be 5×10^{16} for a noon temperature of 360°K, and from (15) a less value for a higher temperature. From the dotted curve of Fig. 1, assuming that the ionization is in terms of pairs of electrons and positive ions as given by the lower scale of abscissas, and that y of F_2 is 10×10^5 constant from 300 to 400 km, y_0' is 3×10^{13} charged particles. This is about 1/1700 of the 5×10^{16} required by the diamagnetic theory. If, however, the portion of the curve from 150 to 200 km is mainly ions as given by the upper scale of Fig. 1, the 5×10^{16} charged particles can be accounted for.

No clear explanation of the observed minimum of N at about 10 A.M. at the equator, Fig. 2, is offered by the diamagnetic theory; the simple expression (17) of course gives a minimum at noon. It was pointed out[15] that an increase of the average height of the ionization during the day of 30 km to a maximum at 2 P.M. would advance the phase of the diamagnetic effect into the forenoon hours. Ionosphere height measurements[7] indicate a maximum height at about noon but do not definitely rule out the possibility of a two hour shift into the afternoon. This is due not so much to experimental inaccuracy as it is to the fact that the measurement always yields a virtual or optically equivalent height and not a true height.

In a discussion of diamagnetism Cowling[24] concluded that there exist electric currents at the boundary of a diamagnetic region which cancel the effect of diamagnetism outside of the region. The argument was a repetition of an older one by Bohr (thesis) and by Van Leeuwen[25] in connection with the theory of diamagnetism of metals. In reaching the conclusion Cowling assumed implicitly that the collisions of ions were perfectly elastic. If, however, it is assumed, and there is experimental evidence in support of the assumption,[26] that after collision there is as good a chance of an electron or ion starting out in one direction as in any other, then the peripheral currents disappear and the diamagnetic effects remain.[11] Thus, Cowling's analysis points out how dependent the phenomena are on the exact mechanism of collision.

Altogether the diamagnetic theory is in about the same position as the dynamo theory in that it requires a number of charged particles in the high atmosphere larger than the number at first sight indicated by ionospheric measurements but the existence of which is not definitely denied by the measurements. It is simpler than the dynamo theory since it does not call for a world-wide wind system in the high atmosphere.

Drift current theory

A qualitative theory of S was sketched out by Chapman[13] based on gravitational magnetic drift currents of the long free path ionization. For the more general case in which H and g make an angle ϕ with each other (11) becomes

$$v = mg \sin \phi / He. \tag{19}$$

For a uniformly magnetized sphere, and the earth approaches this quite closely, H at latitude θ is given by

$$H = H_0(1 + 3 \sin^2 \theta)^{\frac{1}{2}}, \tag{20}$$

[23] Chapman, Terr. Mag. and Atmos. Elec. **34**, 1 (1929).

[24] Cowling, Roy. Ast. Soc., Monthly Notices **90**, 140 (1929); **92**, 407 (1932).

[25] Van Leeuwen, J. de physique **2**, 361 (1921).

[26] Darrow, *Electrical Phenomena in Gases* (1932), p. 200.

where $\cos\theta = 2\tan\phi$ and $H=0.32$ gauss is the value of H at the equator at sea level. From (19) and (20)

$$v=\frac{mg}{He}\frac{\cos\theta}{1+3\sin^2\theta}, \tag{21}$$

from which $v=0$ at the magnetic poles and $v=8.5$ cm sec.$^{-1}$ at the equator.

If the total number y' of ions in a 1 cm^2 vertical column of the long free path region is made up of equal numbers of positive and negative ions of equal mass and charge the eastward gravitational magnetic drift current is

$$i=vey'. \tag{22}$$

The eastward drift of the positive ions causes an accumulation of positive charge over the afternoon hemisphere; similarly, negative ions drift westward and accumulate over the morning hemisphere. Thus, a westward e.m.f. is set up in the long free path region. The region, however, is electrically connected to the underlying short free path region for charges can move freely along H to the short free path region below, particularly in the higher latitudes where H is inclined at a considerable angle to the vertical. Therefore the westward e.m.f. causes currents in the short free path region and, if there were a suitable world-wide distribution of conductivity in the short free path region, current systems similar to those of Fig. 3 would be developed.

Putting the equatorial noon value of $i=3\times10^{-5}$ from Fig. 3 into (22) gives $y'=2\times10^{14}$, which is less than the number required by the diamagnetic theory. The drift current theory offered no explanation of the forenoon maximum of S. At this point Chapman left the theory with the suggestion that perhaps dynamo action could somehow be called upon to account for the advance in phase.

The long free path ionization may contribute to S through its diamagnetism and its drift currents. In order that the first be effective the distribution over the earth must be given approximately by (18); for the second to be effective a similar distribution of long free path ionization is required and in addition a proper (but unspecified) distribution of short free path ionization or conductivity. A quantitative investigation[9] was carried out to determine what the distributions might be, the results were favorable to the diamagnetic theory and unfavorable to the drift current theory.

In the investigation the following assumptions were made: (1) that the ionization was caused by the ultraviolet light of the sun; (2) that in the long free path region directly underneath the sun $y'=5\times10^{16}$ in accord with the diamagnetic theory, the ionization being, for example, in a bank from 150 to 200 km with $y=5\times10^9$; and (3) that at the subsolar point in the short free path region y was 5×10^9 at 150 km and decreased approximately exponentially to 1.4×10^7 at 100 km, which led to $\sigma=1.44\times10^{-5}$. In Fig. 1 the shaded portions indicate the nature of the last two assumptions. The shaded area should not extend outside of the dotted curve for that would mean an assumed ionization greater than observed. However, there is no discrepancy for the dotted curve refers to sunspot minimum and the shaded area to an epoque halfway between sunspot minimum and maximum; the ionization is known[27] to have increased by roughly 50 percent from sunspot minimum in 1923 to maximum in 1928.

From the three assumptions the ionization over the earth was worked out taking into account ionic recombination, diffusion and electric and magnetic drift; the details were complicated, the original papers[9] should be consulted for them. It came out that the long free path ionization was in accord with (18) and hence in accord with the diamagnetic theory. The daytime conductivity of the short free path region was high in tropical latitudes and fell to relatively low values at latitudes above 40°; in the night hemisphere the tropical conductivity was about 1/5 the day value. Therefore the electrical circuit of the eastward drift currents of the long free path ions was completed not by currents sweeping westward in high day latitudes as Fig. 3 would have it, but partly (3/4) by a westward current in the day tropical short free path region and partly (1/4) by an eastward current in the night tropical short free path region.

In more detail, the world-wide current system consisted of a flow mainly along the parallels of

[27] Hulburt, Physics **4**, 196 (1933); Young and Hulburt, Phys. Rev. **50**, 45 (1936).

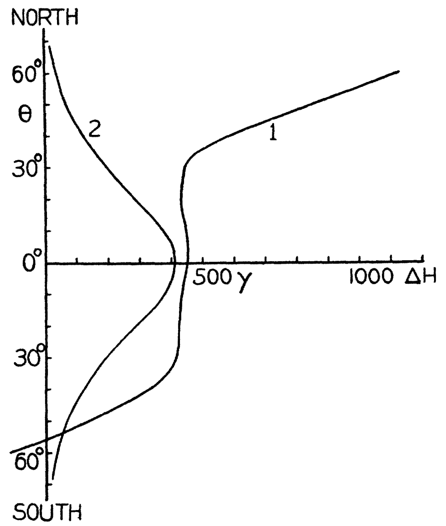

FIG. 4. Curve 1 gives the observed,[28] curve 2 the theoretical[9] values of ΔH, the horizontal component of the portion of the earth's permanent magnetic field of external origin.

latitude in the following way: (1) a current sheet in the daylight hemisphere flowing eastward in the long free path region which at the sunrise and sunset longitudes divides into two sheets; (2) one of these flows westward on the day side of the earth in the short free path underneath (1); and (3) the other sheet continues eastward around on the night side of the earth. The current is mainly (4/5) between the 40th parallels of latitude north and south, and falls to lower values at the higher latitudes. The total currents in the three sheets are about 2.1×10^7, 1.6×10^7 and 5×10^6 amperes, respectively. The east and west daytime current sheets subtract from each other leaving in effect an eastward current of about 5×10^6 amperes flowing around the earth all the time. The distribution of the current with latitude has the shape of curve 2, Fig. 4.

From a harmonic analysis of the 1922 survey of the permanent magnetic field of the earth Bauer[28] concluded that a portion, about 2 percent, is of external origin. His values of the horizontal component ΔH of this portion for various latitudes are plotted in curve 1, Fig. 4; the scale of abscissas is in γ or 10^{-5} gauss. The calculated horizontal magnetic field of the 5×10^6 ampere current flowing around the earth is given in curve 2, Fig. 4. The calculated and observed curves are in fair agreement in tropical latitudes. At high latitudes there is some disparity, but at high latitudes the observed curve is uncertain for the polar data on which it was based were meager. The calculated and observed values of the vertical component of the magnetic field, which are not given here, show an agreement similar to that of the curves of Fig. 4.

The conclusion was reached that the gravitational magnetic drift currents, instead of producing the current sheets of Fig. 3, gave rise to a current encircling the earth of the type to account for the portion of the permanent magnetic field of the earth of external origin. Analysis[9] showed that the conclusion did not depend critically on the exact form of the assumed ion banks in the shaded area of Fig. 1.

We bring this section to a close with the rather trivial remark that a correct explanation of S may involve to varying degree the three effects in the high atmosphere, dynamo action, diamagnetism and gravitational magnetic drift currents.

SECTION C. LUNAR VARIATION OF TERRESTRIAL MAGNETISM

Observational data

A small variation L, about 3γ in amplitude, of the terrestrial magnetic field is recognized as being due to the moon. The character of L is shown in Figs. 5 and 6. In the upper eight curves of Fig. 5 the east component of L in the horizontal plane is plotted against lunar hours from observations at Batavia[29] for various phases of the moon. It is seen that the variation is much greater during the solar day than at night. An average of the eight curves, given in the lower curve of Fig. 5, eliminates the effects of the sun on L and shows that L is of symmetrical semidiurnal type similar to a tidal variation. In Fig. 6 are given the curves of N, V and W, the north horizontal, vertical and west horizontal components of L, respectively, for various latitudes averaged from observations at a number of stations.[30]

As in the case of S harmonic analysis indicated

[28] Bauer, Terr. Mag. and Atmos. Elec. **28**, 1 (1923).

[29] Bartels, *Die Höchsten Atmosphärenschicten* (Berlin, 1928).

[30] Chapman, *Dictionary of Applied Physics*, Glazebrook **2**, 543 (1922).

that three-fourths, or two-thirds, of L arises from overhead influences and one-fourth, or a third, from induced currents in the earth. If L is attributed to the magnetic effects of overhead current systems the maximum current density i at the equator is about 4×10^{-6} e.m.u. The amplitude of L was found to vary with lunar distance nearly in the same ratio as the moon's tide producing force. In contrast with S, L varies but little with annual mean sunspottedness, although at the same epoch L varies greatly from day to day with the magnetic activity.

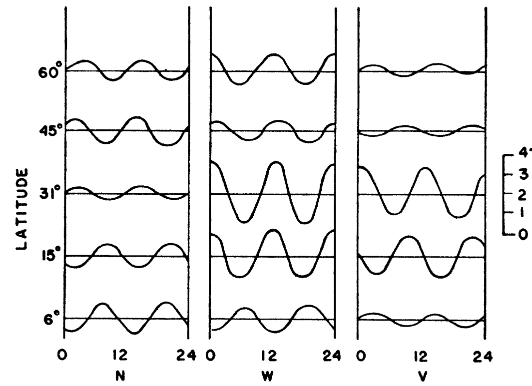

FIG. 6. Average lunar variation of terrestrial magnetic elements.[30]

Dynamo theory

There is only one theory of L, the Schuster-Stewart dynamo theory, which was developed in detail first by van Bemmelen[31] and later by Chapman.[22] The theory ascribed L, just as in the case of S, to the magnetic effects of electrical currents in the high atmosphere. The currents were assumed to be impelled by electromotive forces due to lunar tidal movements of conducting regions of the high atmosphere across the vertical component of the terrestrial magnetic field, the tidal motion being calculated from the lunar barometric variation. Barometric data at Batavia, near the equator, indicated an amplitude of 0.063 mm of mercury for the 12 hour (lunar hour) lunar barometric component; its variation with latitude has not been determined. It was found that the electrical current systems derived from the calculated atmospheric tidal movements were of the type to account for L. There emerged, however, a phase difference of 233° between the theoretical and observed values of L, the two being nearly opposite in phase. This was, and is, difficult to explain, for the lunar barometric variation can hardly be other than of tidal origin.[20] In this respect the L theory is less open to question than the S theory which at all times must envisage the complicating possibility of both thermal and tidal effects. The discrepancy in phase of the L theory was, and is, correspondingly more acute than the similar discrepancy in the S theory.

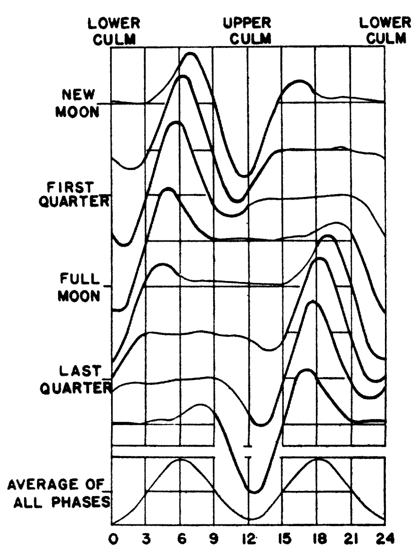

FIG. 5. Lunar variation in $-W$ at Batavia.[29]

To determine the high atmospheric conductivity σ from L we present, as was done for S, a simple calculation instead of the more complete one.[22] From the lunar barometric amplitude 0.063 mm the average velocity v of atmospheric tidal ebb and flow is $1/0.063\times760$ of the peripheral velocity of the rotation of the earth. Hence $v=2$ cm sec.$^{-1}$ or 0.07 km hr.$^{-1}$. Putting $i=4\times10^{-6}$, $H=0.3$ gauss and $v=2$ in $\sigma=i/vH$ yields $\sigma=5\times10^{-6}$. This, however, is an average value of σ around the earth at the equator, since L was derived from an average over a lunar month as indicated in the lower curve of Fig. 5. Hence the noon maximum value of σ is greater by a numerical factor 3 to 5, depending on the assumed law of the diurnal

[31] van Bemmelen, Meteor. Zeits. **5**, 218 (1912); **12**, 589 (1913).

variation of σ. Assuming that σ is a function of the cosine of the zenith angle of the sun we multiply by 4 to obtain an equatorial noon value, and obtain

$$\sigma = 20 \times 10^{-6}.$$

Approximately the same value was reached by Chapman; it is about 7 times the value obtained by Schuster from S. Depending as it does on a theory only partially successful its correctness is uncertain.

Qualitative suggestions were made toward elucidating the departures of the dynamo L and S theories from observation, such as atmospheric resonance or viscosity and the possibility that the S and L actions occur in different levels of the atmosphere.[22] No set of hypotheses which reconcile the various discrepancies has yet been outlined. Further advance in the theories along the lines followed by previous investigators does not seem immediate, for this would entail statistical extraction of very small quantities from long series of data of many observatories scattered over the earth.

Section D. Magnetic Storms and Aurorae

Data of magnetic storms

The erratic variations or disturbances of the terrestrial magnetic field are so irregular as to defy complete classification. They may occur at any time, they may be world-wide or local. The large disturbances, classed as storms, are almost always world-wide and are usually accompanied by auroral displays. Of the large disturbances those with a sudden commencement are recognized as a definite type. The curves of the continuously registering instruments of the magnetic observatory may have been quiet for some hours, nothing in their appearance suggesting a termination of quiet conditions. Then suddenly a sharp movement begins which in low and middle latitudes is normally much largest in the horizontal component N. In a few minutes N may have risen 50γ or more. A considerable proportion of the largest storms are preceded by these movements and they are not unnaturally regarded as precursors or commencements of the storm. In Fig. 7 are given the traces of N recorded at Cheltenham Observatory, Maryland, for three average world-wide storms,[32] illustrating a sudden commencement of the type mentioned in which N increased rapidly, a storm which began with a rapid rise and fall of N, and a storm which set in gradually. Curves of N at a number of observatories[33] are shown in Fig. 8 for the short well-defined storm which commenced suddenly at about 7 hours G.M.T. March 14, 1922.

In some cases of world-wide sudden commencements the initial pulse is immediately followed by highly disturbed conditions, in other cases some hours elapse before any further large movement occurs, and in other cases nothing follows during the next 24 hours even remotely resembling a storm. The world-wide sudden commencement appears at all stations simultaneously within 3 minutes.[33, 34] In Fig. 9 is shown the beginning of the magnetic storm of May 13, 1921, which commenced at 13.10 G.M.T.[33] It was one of the two greatest magnetic storms that occurred between 1908 and 1928. It continued for five days and during that time there were four impulses of the sudden commencement type each one of which was simultaneous over the entire earth.

Aside from sudden commencements of magnetic storms there are almost no prominent features that persist completely around the earth. An outstanding peak, or a sharp depression, found on a magnetogram almost invariably fades out before it gets halfway around.[33] In examining the magnetograms of twelve years, 1913 to 1924, Wallis[33] found one short sharp disturbance that could be definitely traced entirely around the earth. This was the sharp wedge-shaped depression in the curves of Fig. 9 near 15.15 G.M.T.

In the records from tropical[35] and middle latitudes any kink in the curve of the characteristic appearance of the sudden commencement is almost always world-wide. This is not true for records from higher latitudes. For example, at Eskdalemuir, Scotland, a considerable number of sudden movements may be found in an average

[32] Cheltenham Observatory Records (1928).
[33] Wallis, Terr. Mag. and Atmos. Elec. **35**, 93 (1930); **36**, 15 (1931).
[34] Bauer, Terr. Mag. and Atmos. Elec. **15**, 9, 221 (1910); Faris, ibid. **15**, 93 (1910); Rodes, ibid. **27**, 161 (1922).
[35] Moos, Bombay Magnetic Observations, 1846–1905.

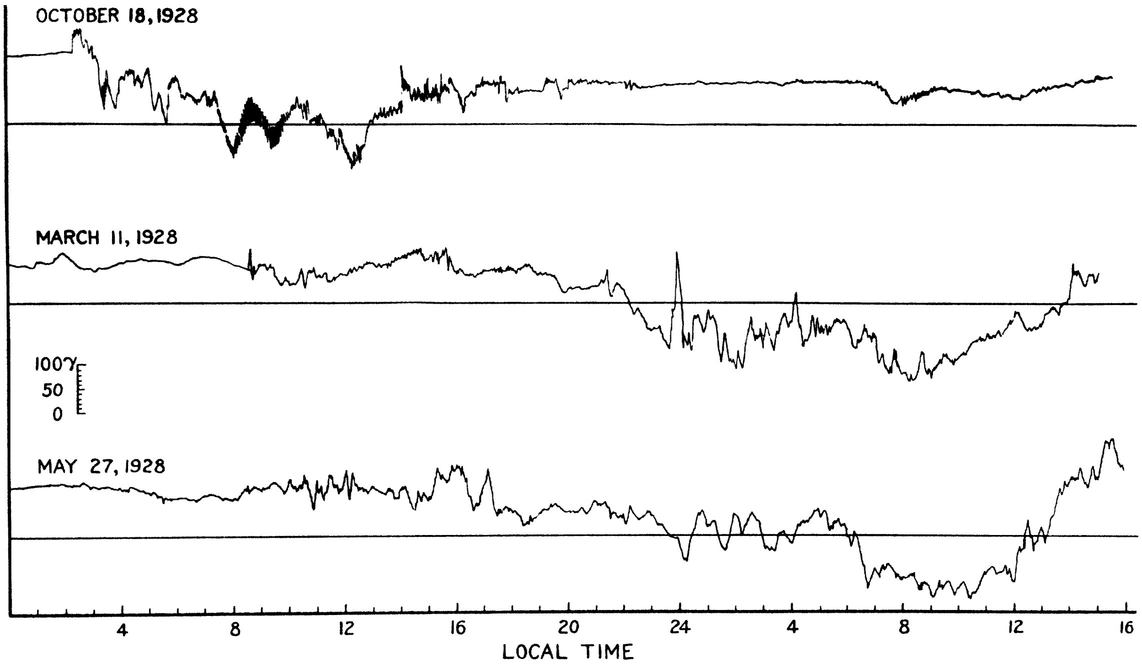

FIG. 7. Horizontal component of three world-wide storms at Cheltenham Magnetic Observatory, Maryland.[32]

year in addition to the world-wide ones which are not represented at distant stations.[36] At polar stations the record is so agitated during disturbance periods that almost any feature sought for can be found.

The initial enhancement of N in the sudden commencement is often referred to as the first phase or the impetus[37] of the storm. Sooner or later this is often, but not always, followed by a marked depression of N, referred to as the second phase, in which though generally interrupted by irregularities and oscillations N falls well below the prestorm value descending perhaps 100γ in several hours. Finally a recovery period ensues which in turn may be interrupted by oscillations. Some of these facts are illustrated by the curves of Figs. 7 and 8. Chapman[38] selected about 40 storm curves of this type from a number of observatories, averaged them and after subtracting the quiet day diurnal variation separated them into two features, those features regarded as world-wide and hence measured in "storm time" or hours from the beginning of the storm, and those features characteristic of the local time of the observatory or the diurnal storm variations. The average curves[38] are given in Figs. 10 and 11. It is seen from Fig. 10 that the world-wide storm variations are about the same at high and low latitudes, and from Fig. 11 that the diurnal storm variations increase in intensity with latitude.

The energy density E of a medium of permeability μ in a magnetic field H is

$$E=\mu H^2/8\pi. \tag{23}$$

Let x_0, y_0, z_0, and x, y, z be the mean undisturbed and disturbed components of H, respectively, along rectangular axes, and $\Delta x=x_0-x$, $\Delta y=y_0-y$, $\Delta z=z_0-z$. Then the change in E during disturbance is, for $\mu=1$,

$$\Delta E=(x_0\Delta x+y_0\Delta y+z_0\Delta z)/4\pi+(\overline{\Delta x^2}+\overline{\Delta y^2}+\overline{\Delta z^2})/8\pi, \tag{24}$$

or to a close approximation, neglecting the second term,

$$\Delta E=H_0\Delta H/4\pi. \tag{25}$$

[36] Chree, British (Terra Nova) Antarctic Expedition (1910–1913); Terrestrial Magnetism (1921).

[37] Angenheister, Nachr. d. ges. d. Wiss. Göttingen, S1 (1924).

[38] Chapman, Proc. Roy. Soc. **A95**, 61 (1919); **A115**, 242 (1927); Terr. Mag. and Atmos. Elec. **40**, 349 (1935).

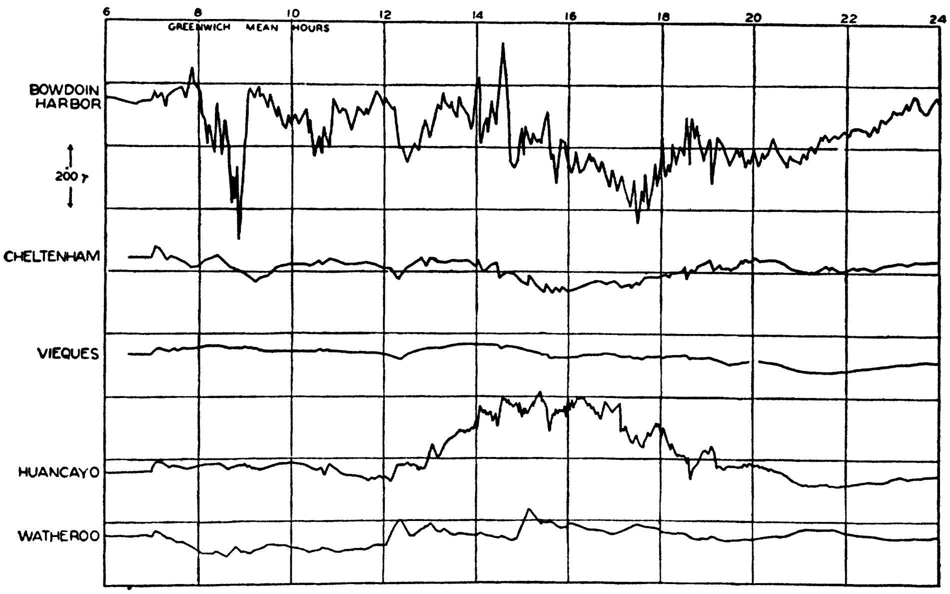

FIG. 8. Horizontal component at different stations for the magnetic storm of March 14, 1922.[33]

Wallis[33] calculated the average ΔE for the storms of March 14, 1922, and January 29, 1924, of duration 18 and 26 hours, respectively, using data from the magnetic observatories at Bowdoin Harbor, Sodankyla, Sitka, Cheltenham, Tuscon, Viecques, Honolulu, Antipolo, Huancayo, Pilar, Vassouras and Watheroo. The results, plotted in Fig. 12, show that ΔE passes through a broad minor maximum near the magnetic equator, a shallow minimum between magnetic latitudes 20° and 40°, rises to a high maximum between magnetic latitudes 60° and 80°, and descends less steeply toward the magnetic poles, the disturbance energy at the poles being about the same as at 60°. The march of the ΔE curve agreed with the curve of average auroral frequency.[39] Identical conclusions were reached by Stagg[40] in a similar investigation.

From Fig. 12 the average energy density of the storm magnetic field is about 10^{-5} erg cm^{-3}. Multiplying this by the volume of the earth gives 10^{22} erg as the total energy of the storm, which may be increased to about 3×10^{22} to account for the storm field in space. If this energy were received from the sun in 28 hours, the rate of reception is 3×10^{17} erg sec.$^{-1}$ which is only a small fraction of the 2×10^{24} ergs of sunlight intercepted by the earth each second.

Magnetic disturbance, radio, ionosphere and sun

With the establishment of long distance communication on short radio waves 15 to 40 meters in length it was early observed that the circuits were disturbed and often rendered inoperative during strong magnetic disturbance. Analysis[41] showed that the perturbation was essentially a daylight effect as follows: daylight circuits were disturbed at the commencement of the storm, the full night circuits remaining normal until dawn when they were, or might be, disturbed; the disturbance in the daytime circuits often persisted after nightfall.

Detailed knowledge of the ionosphere during magnetic disturbance is as yet fragmentary. Schafer and Goodall[7] concluded from their observations and those of others that during severe storms there was an increase in the absorption of the radio echo together with a scattering or splitting which manifested itself by the large number of reflections returned at slightly different virtual heights and which rendered it difficult to distinguish one ionospheric layer from another. During moderate storms it

[39] Fritz, *Das Polarlicht* (1881); Boller, Gerlands Beitr. **3**, 56, 550 (1908).
[40] Stagg, Terr. Mag. and Atmos. Elec. **40**, 255 (1935).
[41] Maris and Hulburt, Proc. Inst. Rad. Eng. **17**, 494 (1929).

was usually difficult to tell whether a storm was in progress from qualitative observations of the reflections. f_c of E and F_2 did not change very much with magnetic character, f_c of F_1 became less with increase of magnetic character, but rarely by more than 20 percent.

The program of world-wide ionospheric research is being expanded rapidly with continuous recording equipment of increased power and range of radio spectrum. It seems certain that in the next few years many results will emerge of fundamental importance to the elucidation of terrestrial magnetism.

Even now new discoveries have been announced which suggest the course of future progress. Dellinger[42] has called attention to a new phenomenon of short wave radio transmission which occurs on the illuminated half of the globe. It is a sudden disappearance of radio signals for a few minutes the complete process of fading out and reappearing occupying about 15 minutes; it is simultaneous on all full daylight circuits and is absent from nocturnal circuits. The radio fadeouts occurred during periods of magnetic disturbance, and during 1935 were spaced at intervals of about 54 days. On the idea that some unusual form of sunspot activity might be responsible for the fadeouts the hydrogen H_α spectroheliograms made at Mount Wilson Observatory were examined.[43] On two fadeout dates, July 6 and August 20, 1935, spectroheliograms were available at the correct times, and showed sudden marked changes in the form and intensity of a hydrogen flocculus within a few minutes of the time of the fadeout. No complete ionospheric data were taken on the two dates.

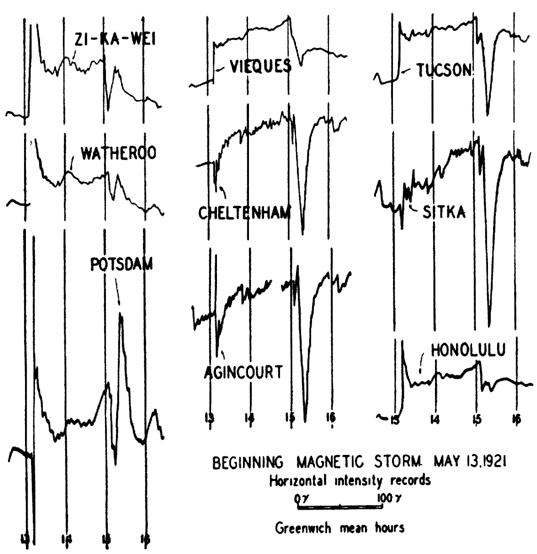

FIG. 9.

Two things were evident, that the phenomenon was short-lived lasting only a few minutes, and that its further elucidation would call for a continuous watch of the sun as well as of radio circuits and the ionosphere; continuous recording of terrestrial magnetic variations has been *un fait accompli* for many years. The recurrence period of about 54 days, although its reality was not stressed, at least indicated when such a watch might best repay the effort.

On October 24, 1935, a moderately bright hydrogen eruption occurred together with a partial fadeout involving the loss of the higher radiofrequencies; there was a world-wide magnetic storm. Ionosphere observations[44] showed that f_c of F_2 increased from October 10 to 23 to a very high value, then on October 24, dropped to one-half, and on October 25, and succeeding days returned to the previous high value. At the same time the virtual height of F_2 shot up to 460 km on October 24 from a height of about 250 km on the preceding and following days.

Following this, bright H_α eruptions were recorded on December 16 and 17, 1935, but no very unusual radio phenomena were reported near these dates. Fadeouts occurred on February 6, 14, and 17, 1936, the one on the 14th being very pronounced. No unusual solar activity was noted on February 6 and cloudy weather prevented observation at Mount Wilson from February 10 to 19.

On April 8, 1936, an exceptionally brilliant solar eruption was observed at Mount Wilson[43] and at Huancayo Peru,[45] which began at 16^h 45^m G.C.T., increased in brightness until 16^h 47^m and returned to normal at about 17^h 03^m. A striking and wide-spread fadeout of daylight high frequency radio circuits began at 16^h 46^m G.C.T., April 8, lasting from 15 to 30 minutes

[42] Dellinger, Phys. Rev. **48**, 705 (1935); Science **82**, 351, 548 (1935).

[43] Richardson, Trans. Amer. Geophys. Union (1936).

[44] Kirby, Gilliland, Judson and Smith, Phys. Rev. **48**, 849 (1935).

[45] Torreson, Scott and Stanton, Science **83**, 463 (1936).

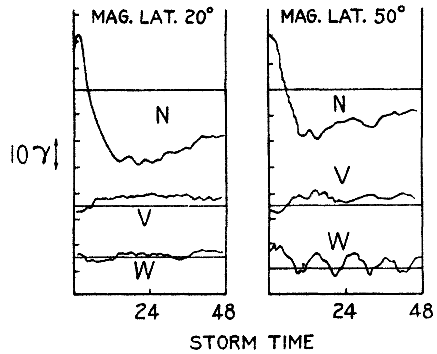

FIG. 10. World-wide storm variations at various latitudes.[38]

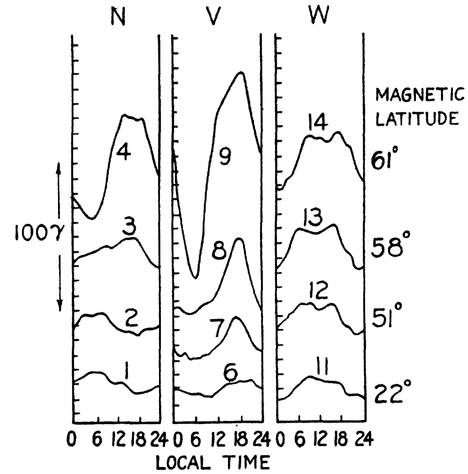

FIG. 11. Diurnal storm variations at various latitudes.[38]

for various radiofrequencies. A world-wide magnetic storm commenced suddenly at $16^h\,46^m$; at Huancayo N increased 108γ from $16^h\,46^m$ to $16^h\,51^m$. Radio echoes from the ionosphere at Huancayo suddenly ceased at $16^h\,45^m$ and returned again to normal in a little over an hour. The bright hydrogen eruption, the radio fadeout and the sudden change in N and the ionosphere were all phenomena of an unusual and definite character. Richardson[43] concluded that in this case at least there was strong evidence that the solar activity was directly connected with the terrestrial phenomena and that the energy was transmitted with the velocity of light.

Data of aurorae

Probably the most complete summaries of the facts of the aurora are those of Chree[46] and of Vegard.[47] There is a zone of maximum auroral frequency[39] about 23° from each magnetic pole as shown in Fig. 13 in which the average number of auroral displays in a year is plotted in curve 1 against magnetic latitude. At latitudes above the 67° maximum, curve 1 is dotted for no data are given, merely the statement that in high latitudes aurorae are less frequent. Diffuse auroral clouds and pulsations are observed at heights from about 85 to 150 km and rays and streamers from about 110 to 600 km.[48, 49]

Near the maximum frequency zone aurorae are of almost daily occurrence; they spread into lower latitudes during magnetic disturbance, the spread being greater for the more intense storms. During relatively calm magnetic conditions there appears to be no uniformity in the diurnal course of auroral display at various stations, the data at any station giving the impression that the aurora depended on the local meteorology of the high atmosphere above the station such as winds and conditions of excitation of the atmospheric atoms and molecules.[50, 51] During magnetic storms intense, moving and colored displays are usually more frequent in the evening than in the morning, the diffuse, weak and quiet forms being more abundant in the morning.

Many of the lines in the spectrum of the aurora have been identified[52] as emission lines from nitrogen molecules and oxygen atoms, the prominent green line at 5577A originating from metastable atomic oxygen. No helium or hydrogen emissions have been detected. Störmer[53] has called attention to very high auroral streamers of a violet gray color extending occasionally to an altitude of 1000 km. These were in a region illuminated by the sun. Their spectra differed

[46] Chree, *Encyclopaedia Britannica*, thirteenth edition **2**, 927 (1926).
[47] Vegard, *Handbuch der Exper. Physik* **25**, 1, 385 (1928).
[48] Vegard, Phil. Mag. **42**, 47 (1921); Vegard and Krogness, Geophys. Pub. Oslo, Nr. 1 (1920); Störmer, ibid., Nr. 7 (1926).
[49] Curie, Terr. Mag. and Atmos. Elec. **39**, 293 (1934).
[50] Hulburt, Terr. Mag. and Atmos. Elec. **36**, 23 (1931), summarized the data from many expeditions in the north and south.
[51] Fuller, Terr. Mag. and Atmos. Elec. **40**, 269 (1935).
[52] McLennan, Proc. Roy. Soc. **A120**, 327 (1928); McLennan and collaborators, Proc. Roy. Soc. **A115**, 515 (1927); **108**, 501 (1925); **106**, 188 (1924).
[53] Störmer, Zeits. f. Geophys. **5**, 177 (1929).

from those of the usual aurorae in the earth's shadow in that the green line was relatively weaker than the nitrogen bands.[53, 54]

Corpuscular theory of magnetic storms

There is no complete corpuscular theory of magnetic disturbance. At various times charged particles from the sun have been suggested to account for magnetic effects but the suggestions have been either indefinite or shown to be untenable. It has been long recognized[55] that the world-wide increase in N during the first phase of the storm could be occasioned by a sudden existence of an eastward electric current ringing the earth mainly in tropical latitudes, and the decrease in N during the second phase by a similar westward current. Such ring currents formed of electrons or positive ions have been postulated[56] but no investigations were made of the manner of formation, of the stability and disintegration of the rings. The postulates appeared to create more difficulties than they solved.

A corpuscular theory of magnetic storms developed by Chapman[57] did not survive the criticism of Lindemann[58] and was given up by the author.[57] The theory assumed that particles of like charge, emitted by the sun in a beam, plunged into the outer atmosphere causing a downward movement of the conducting region. By dynamo action, Eq. (12), this induced an eastward current around the earth which caused the increase in N of the first phase of the storm. Due to electrostatic repulsion of the charged particles the downward motion was followed by an upward expansion which induced a westward current, the second phase of the storm. However, Lindemann pointed out, making use of an idea advanced by Schuster,[59] that a beam of particles with the same sign of charge and of sufficient density to carry the storm energy could not proceed as a beam for more than one or two solar diameters on account of electrostatic repulsion and could not approach the earth after the first few seconds on account of the charge which the earth would rapidly acquire.

Chapman and Ferraro[60] attempted to construct another theory on the assumption that electrically neutral streams of equal densities of positively and negatively charged particles are shot out from the sun. They envelop the earth at distances of 50,000 km or more and through their diamagnetism cause the increase in N of the first phase of the storm. As the cloud of charged particles progresses into the terrestrial magnetic field the positives are diverted westward and the negatives eastward. The idea was entertained that the particles might swing around the earth and form a westward electrical current which would cause the second phase of the storm. However, the authors concluded that they were unable to follow through the physics of the establishment of the ring current. Chapman recently remarked,[61] "The whole theory is neccessarily both speculative and difficult; probably the most doubtful feature is that relating to the ring current, the existence and formation of which are still very uncertain." The theory was left in a qualitative and unfinished state. It suggested no explanation of the diurnal storm variations, nor did it envisage how the particles were to enter the ionosphere to produce changes during magnetic disturbance.

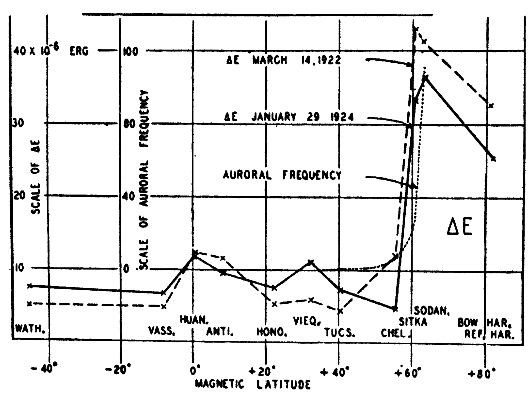

FIG. 12. Mean values of energy density ΔE of magnetic storms of March 14, 1922, and January 29, 1924.[33]

[54] Kosirev and Eropkin, Poulkovo Obs. Circular, No. 18 (1936).

[55] van Bemmelen, Terr. Mag. and Atmos. Elec. **8**, 153 (1903); Schmidt, Met. Zeits. **34**, 385 (1899).

[56] Störmer, C. R. Acad, Sci. **131**, 736 (1910); Birkeland, Norwegian Aurora Polaris Expedition (1902–1903).

[57] Chapman, Proc. Roy. Soc. **A95**, 61 (1918); **A115**, 242 (1927); Proc. Camb. Phil. Soc. **21**, 577 (1923).

[58] Lindemann, Phil. Mag. **38**, 669 (1919).

[59] Schuster, Proc. Roy. Soc. **A85**, 44 (1911).

[60] Chapman and Ferraro, Terr. Mag. and Atmos. Elec. **36**, 77, 171 (1931); **37**, 147, 421 (1932); **38**, 79 (1933).

[61] Chapman, *Quatrième Rapport de la Commission pour l'Étude des Relations entre les Phénomènes Solaires et Terrestres* (1936), p. 56.

Corpuscular theory of aurorae*

The well-known corpuscular theories of the aurora of Birkeland, Störmer and Vegard[47] are confronted with the Schusterian[59] difficulty which they have never got around, as pointed out by Lindemann,[58] Swann[62] and others. Störmer's calculations of the focusing of the charged particles into the auroral zones by the magnetic field of the earth refer to a single particle, as an electron or an α-particle, whereas the auroral theory requires a cloud of particles, and it has been shown[58, 60, 62] that a stream of electrons or ions, or mixtures of electrons and ions cannot account for the auroral zone. Chapman and Ferraro[63] concluded, "A reexamination of the conditions of passage of the stream from the sun to the earth disposed of our lingering hope that the stream might carry some small residue of charge which would at least suffice to explain the production of aurorae, by permitting the stream to be deflected in the same way as separate corpuscles are in Störmer's theory (although to a smaller extent); our work confirmed Lindemann's conclusion that the only admissible kind of stream is one that is electrostatically neutral to a very high degree of approximation."

Ultraviolet light theory, general assumptions

The ultraviolet light theory of magnetic storms and aurora was developed in detail, largely quantitative, in a series of papers.[41, 50, 61] The more important physical ideas are described here. Table I shows that above a level of about 300 km there are approximately 10^{16} molecules in a 1 cm^2 vertical column of the atmosphere. Above this level they experience practically no collisions but dance up and down receiving upward thrusts from impacts from below and falling back under gravity. This is the outer fringe or spray of the atmosphere.

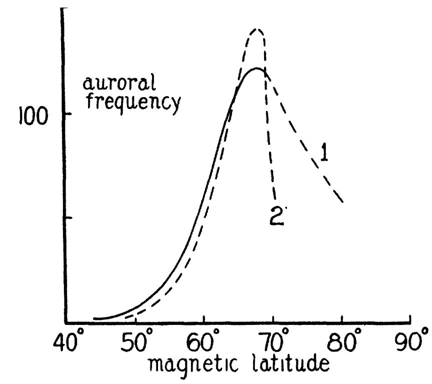

FIG. 13. Average auroral frequency, curve 1 observed,[39] curve 2 theoretical.

The 10^{16} molecules experience 10^{14} collisions sec.$^{-1}$ at the 300 or 400 km level. For a temperature of 360° Kelvin ordinary impacts impart velocities of order 1 km sec.$^{-1}$ which rarely send the particles above 2000 km. It is assumed that during solar quiescence; i.e., no magnetic storms or aurorae, 10^8 of the 10^{14} impacts are collisions of the second kind with excited atoms and molecules which give the particles velocities as great as 10 km sec.$^{-1}$. They reach heights of 40,000 to 80,000 km in 3 to 6 hours and if unionized fall back to the earth in various orbits depending on the angles of projection, as shown in Fig. 14.

It is assumed that the high flying particles are ionized by the ultraviolet light of the sun in 3 to 6 hours. Once ionized they follow along the lines of magnetic force into the outer atmosphere of high latitudes and give up their energy in the form of auroral light and magnetic effects. The flux of energy into various latitudes was calculated and is plotted in the dotted curve 2, Fig. 13; it is in agreement with the observed curve 1. Under the assumptions just outlined the energy transported into the auroral zones is 5×10^{13} erg sec.$^{-1}$ which is perhaps sufficient for a quiet aurora but not as great as the 10^{15} erg sec.$^{-1}$ of a strong auroral display.[64]

It is assumed that the sun when active sends out a blast of ultraviolet light for half an hour. For example, if 1/10,000 part of the solar surface were to emit black-body radiations at temperature 30,000° Kelvin, the solar constant would be increased by 0.74 percent and the solar energy at wave-lengths 3500, 4000, 5000 and 6000A by

* This section is made brief and incomplete as I am informed by the editor that a report on aurorae is being prepared for this journal in which the corpuscular theory will be described.

[62] Swann, Phil. Mag. **47**, 306 (1924).

[63] Chapman and Ferraro, Terr. Mag. and Atmos. Elec. **36**, 80 (1931); Mon. Not. Roy. Ast. Soc. **89**, 470 (1929); Quart. J. Roy. Met. Soc. **52**, 225 (1926).

[64] Hulburt, Phys. Rev. **31**, 1038 (1929); **34**, 344 (1929); **36**, 1560 (1930); Maris and Hulburt, Phys. Rev. **33**, 412 (1929).

3.2, 1.7, 0.75 and 0.32 percent, respectively. Corresponding values which are observed[65] in the short interval (an hour or so) variations of the solar energy are roughly 1, 8, 2, 1 and 0.2 percent, and often more. The energy in the ultraviolet regions 3000–2000, 2000–1000 and 1000–0A would be increased by factors of 1.5, 10^2 and 10^5, respectively. Calculations such as the foregoing based on the assumption of black-body conditions can only be regarded as illustrative approximations, and perhaps very distant ones, for line emission and absorption undoubtedly play a predominating role. The energy increases the ionization of the high atmosphere and the high flying spray and causes heating, outward expansion and winds. These in turn give rise to aurorae and magnetic disturbances.

Ultraviolet light theory of aurorae

In the spectrum of the ultraviolet solar flare the wave-lengths shorter than the limit of the principal series of a given atom are probably the most effective in ionizing the atom whereas somewhat longer wave-lengths would be expected to cause excitation of the atom. In general various types of solar flares may be assumed, such as (1) flares containing mainly the excitation wave-lengths, (2) flares which contain both the excitation and ionizing wave-lengths in varying intensity, etc. A flare of type 1 would produce many excited atoms; these give rise to high flying atoms or molecules which when ionized fall to the polar regions and cause aurorae. Since such a flare does not produce many ions in temperate latitudes, there will be no strong world-wide magnetic disturbance. Therefore, in this case there may be a strong auroral display in polar latitudes unaccompanied by a magnetic disturbance in temperate latitudes of sufficient intensity to be called a storm. Examples of this were the strong auroral displays seen on the Antarctic continent on April 20 and 27, June 18 to 21, 1908, at Cape Royds[66] on June 3, 1912, and on May 29, June 19 and July 7, 1913, at Cape Denison. The magnetic conditions at the stations were disturbed at these times,[66] but in temperate latitudes no magnetic storms were recorded on or within a few days of these dates. Many brilliant auroral displays with no temperate zone magnetic disturbances were recorded by the Maud Expedition[67] during the years 1922 to 1925 in the Arctic Sea north of Siberia.

A flare of type 2 would, in addition to producing many ions, give rise to many high flying particles which, because of the unusual intensity of the ionizing wave-lengths, would be ionized quickly before they had time to attain great heights. They would not reach polar areas, but would move along the magnetic lines of force of the earth's field to high temperate latitudes in sufficient concentration to give an auroral display. Therefore strong magnetic storms would be expected to be accompanied by aurorae in high temperate latitudes and the stronger the storm the more spreading of the display into lower latitudes there would be. This spreading effect has long been recognized.[68] Temperate zone auroral displays were observed by Barnard[69] at the Yerkes Observatory on about 140 nights in the years 1902 to 1909; magnetic storms accompanied 120 of the displays. On the nights when there were no storms the auroral manifestations were feeble and it is possible that magnetic disturbances may have occurred which were too weak to be listed as storms.

A particle ejected upward from the atmosphere with a velocity around 10 km $sec.^{-1}$ reaches 40,000 km levels in about 3 hours. If it is ionized there by the sunlight it falls to polar regions requiring several hours for the descent, the exact time depending on the pitch of the spirals of its path and on its velocity at the instant of ionization, provided that the only influences which act on the ion are gravity and the magnetic field of the earth. Other agencies may be thought of, such as light pressure, electric fields, etc., but whether these are important cannot be said. After falling to the 300 or 400 km level further descent of the ionization is by diffusion through the atmosphere, and calculation[64] shows that 3 to 6 hours are required to

[65] *Annals of Astrophysical Observatory of the Smithsonian Institution IV*, pp. 17 and 207; Smithsonian Miscellaneous Collections **80**, 10 (1927).

[66] Mawson, British Antarctic Expedition 1908, Trans. Roy. of South Australia **40**, 151 (1916); Australasian Antarctic Expedition 1911–1914, Scientific Reports, Series B. 2, Part 1 (1925); Webb, ibid., Series B. 1, March (1925).

[67] Sverdrup, Carnegie Inst. Pub. No. 6, 175 (1927).

[68] Røstad, Geofysiske Pub. Oslo **5**, 1 (1928).

[69] Barnard, Astrophys. J. **31**, 208 (1910).

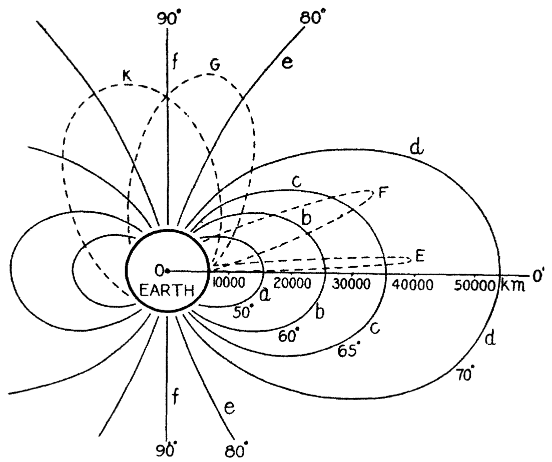

FIG. 14. Curves a, b, c, etc., are the magnetic lines of force of the magnetic field of the earth. The number near each curve is the magnetic latitude where the line touches the earth. Curves E, F, G, and K are the orbits of particles projected from the earth with velocity 10 km sec.$^{-1}$ at angles to the vertical of 1°, 22°, 45°, and 60°, respectively.

reach a 100 km level. If the same ultraviolet flare which causes the polar ion migration and hence the aurora, also causes heavy ionization in the upper atmosphere and hence a magnetic storm, the aurora should appear several hours after the initiation of the storm. Since the ion migration to the poles is, or may be, suppressed while the flare is at its greatest intensity one might expect the delay in the appearance of the aurora in high polar regions to be greater for a long continued intense magnetic storm than for a shorter storm.

The auroral observations recorded in the journals of the British and Australasian Antarctic Expeditions[66, 70] and of the Maud Arctic Expedition[67] furnished data in keeping with the foregoing theoretical inference. In Fig. 15 are plotted the aurora data at Cape Denison, Antarctica, and the world-wide intense (No. 3) magnetic storm which began at 2 A.M. September 17, 1912. The aurora occurred about 40 hours after the storm began. Altogether 11 magnetic storms were found for which the Antarctic and Arctic auroral data were sufficiently complete to permit figures similar to Fig. 15 to be made. The results indicated[64] that the polar auroral displays began about a day after the commencement of the storm in low latitudes, and although the data were too few for great certainty that the time interval between the display and the storm was longer for the more intense storms. Davies[71] in an extended statistical analysis of the auroral observations of the first Byrd Antarctic Expedition presented further confirmation of the delay effect in concluding that, "Comparison of auroral character numbers for days with the international magnetic character numbers shows marked evidence for the occurrence of an auroral maximum on the same day or one day after a maximum in the magnetic character curve."

No reference has thus far been made to the production of the spectrum of the aurora in the laboratory. Kaplan[72] has been able to reproduce many features of the auroral spectrum in weak discharges and after glows in oxygen and nitrogen, and to obtain much information about the excitation processes. In the end the experiments will lead to very important knowledge of the physical state of the atmosphere during aurora.

Finally, it may be remarked that all theories of the aurora are geometrical theories in that they merely offer mechanisms whereby energy is supplied to auroral regions. They make no suggestion concerning the manner in which the energy is transformed into auroral light. A complete theory should do this and will require among other things complete knowledge of the energy levels, metastable states and transition probabilities of the atmospheric atoms and molecules as well as of the exact processes which give rise to the auroral light. Further, the idea has obtruded itself at all times that the emission of the auroral light may not depend simply upon an energy influx carried by ions, but may require in addition that the atmosphere into which the ions fall be in a suitable and perhaps critical condition, such condition being an erratic function of wind currents and states of excitation of the atoms and molecules.

Ultraviolet light theory of magnetic storms

The assumed solar flare of 1/10,000 part of the solar surface at 30,000° would send 10^{22} ergs to the earth in 1 second, which is roughly the

[70] Wright, British (Terra Nova) Antarctic Expedition 1910–1913, "Observations on the Aurora" published 1921.

[71] Davies, Terr. Mag. and Atmos. Elec. **36**, 199 (1931).

[72] Kaplan, Phys. Rev. **45**, 671 (1934); **42**, 86 (1932); **33**, 154 (1929).

total energy of an average magnetic storm of one day duration. Thus a smaller or less intense flare lasting for more than 1 second would supply the storm energy. The light from the flare is mainly in the ionizing wave-lengths below 1500A and increases the ionization of both the long and short free path regions over the day hemisphere; flares with different spectral characteristics may increase the long and short free path ions by different amounts. An increase in the daylight long free path ions increases the eastward 5×10^6 ampere current which girdles the earth N then increases and V decreases as in the first phase of the storm. An increase in the daylight short free path ions causes an increase in the electrical conductivity and hence an increase in the westward daylight current in the short free path ion region. This reduces the eastward 5×10^6 ampere current and therefore amounts to a westward storm current girdling the earth; N is decreased and V is increased as in the second phase of the storm.

There is an additional effect which is common to all types of flares; namely, a heating of the high atmosphere above the, say, 100 km level. The heating causes an outward expansion of the outer atmosphere which produces two reactions, a dynamo effect and an "engulfing" effect. The dynamo effect results from the upward movement across H, and is a maximum in tropical latitudes where H is approximately horizontal. It gives rise to an induced westward electric field and causes a westward current around the earth which reduces N. An upward velocity of the short free path region of Fig. 1 of 10 km hr.^{-1} reduced N by about 100γ at the equator. The engulfing effect, also a maximum in tropical latitudes, depends on the fact that where H is approximately horizontal the long free path ions can only move a short distance across H between collisions and are therefore hindered from moving upward freely with the neutral particles of the expanding atmosphere. Thus, after the initiation of the storm many of the long free path ions find themselves in a mounting tide of air molecules; their free paths are shortened and they become short free path ions. Since the ions experience many collisions before being lost by recombination, the daylight short free path ions are increased at the expense of the long free path ions with a consequent effective westward storm current.

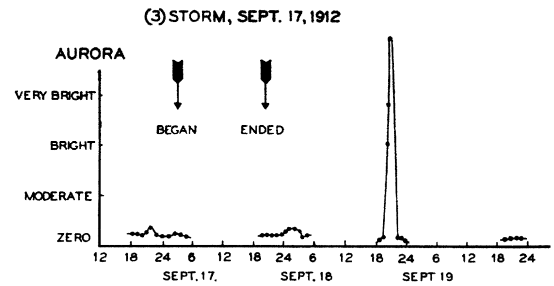

FIG. 15. Aurora intensity at Cape Denison, Antarctica, and magnetic storm of September 17, 1912.

We may picture the sequence of events in the world wide storm of Fig. 10 as follows: the solar flare by its ionizing energy first increases the long free path ions which enhances the eastward current and increases N to a value above normal as in the initial phase of the storm. In the second phase the heating effects of the flare make themselves felt, the high atmosphere expands, the ionized layers move upward and N decreases for a few hours. When the layers cease their upward movement N would increase to its previous value above normal were it not for the fact that in the meantime the engulfing effect has reduced the long free path ions, so that actually N decreases to some value below normal. Thereafter N increases slowly to its normal value as the flare dies away, the storm ionization diminishes and the atmosphere settles down to its quiet day condition.

Magnetic storms which do not begin with an increase in N, and which consist of only the second phase, are fairly common; e.g., the third curve of Fig. 7. In these cases we may suppose that the flare emits mainly energy which heats the high atmosphere and produces mainly short free path ionization. Storms made up only of the first phase are less frequent, they require a flare which increases the long free path ionization without much heating. Irregular and local storm magnetic variations may be attributed to motions of the ionization due to winds, and to effects of induced currents in the earth.

The diurnal magnetic storm changes of Fig. 11 are attributed to the effects of the high flying ions which descend into polar regions. Assuming that the earth as a whole carries no excess charge it was shown[64] that particles sprayed out from the noon meridian of low latitudes fall as ions into the 300 to 400 km levels of high latitudes at 1 to 2 P.M.; they are slowed up and descend by diffusion the next 200 km in 3 to 6 hours. Being long free path ions they set up eastward magnetic gravitational drift currents which, if the conductivity of the high atmosphere varies with the zenith angle of the sun approximately as given by (16), spread into current sheets as sketched in Fig. 16. This current system gives rise to the diurnal storm changes of Fig. 11. The theory was shown to be acceptable quantitatively. At the same time winds in the ionized high atmosphere blow across the twilight zones from sunlit to dark areas and cause strong irregular magnetic agitation. It seems possible that the increase of magnetic disturbance energy with latitude brought out in Fig. 12 may be due mainly to the effect of winds.

The recent observations of the simultaneity of incipience of radio fadeout, ionosphere change, magnetic disturbance and solar eruption[45] appear to offer detailed support of the ultraviolet theory. However, rather than to stress the agreement prematurely it is better to await the results of future experiment. It seems almost certain that the several phenomena may differ in relative degree; for example, some fadeouts or ionosphere changes may occur with no marked magnetic disturbance or visible solar perturbation, and *vice versa*. Therefore, perhaps only in cases of considerable intensity can unmistakable causal relation be inferred between the phenomena. The wave-lengths of the ultraviolet flare which occasion the terrestrial phenomena are shielded by the atmosphere from direct observation, and hence are invisible. Only if these are accompanied by longer wave-length radiations, and there is no compelling reason why they should be, can the flare be seen. Therefore it may turn out that detection of the flare by observation of the solar surface may fail in many cases. Further, the flare may be only a very minute single pinpoint on the solar disk, or a number of even smaller pinpoints.

In this connection an interesting program of observations of solar prominence spectra, initiated in 1928, has been carried on by Perepelkin[73] at Poulkovo Observatory. The ratio i of the intensity of the Balmer hydrogen line H_ϵ to that of the calcium resonance line H was found to vary with magnetic activity W_D as shown in Fig. 17, where W_D was the diurnal declination range at Pavlovsk. The correlation between i and W_D was better than between the relative sunspot number and W_D. This was a new important experimental fact quite apart from any theory. Perepelkin pointed out that unfortunately the new index i referred only to the edge of the solar disk where the prominence spectra were obtainable. The mean value of i was practically the same for various heliographic latitudes but the individual values varied from prominence to prominence, leading to the supposition that the causes of the change in i should not be looked for on the sun as a whole. A theory was outlined which adopted the hypothesis of the existence on the solar surface of spots rich in ultraviolet radiation. In the case of constancy of the relative concentrations of hydrogen and calcium atoms such radiations increase the intensity of H_ϵ more than that of H and hence increase i.

Beam and flare theories of solar outbursts

Two theories concerning the form of the emission from the sun which gives rise to magnetic disturbance have been advanced. One assumes that the solar emission is in a narrow beam,[74] as from a searchlight or a shotgun, the other that the emission is in a wide flare[64] as from a volcano.

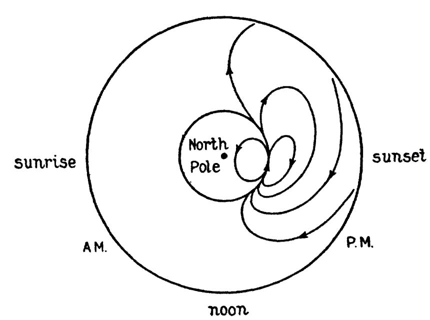

FIG. 16. Theoretical currents in the high atmosphere which cause the diurnal magnetic storm variations.

[73] Perepelkin, Poulkovo Observatory Circular, No. 9 (1933); Gussev, ibid., No. 10 (1934).

[74] Maunder, Mon. Not. Roy. Ast. Soc. **76**, 63 (1913); **65**, 2, 666 (1904).

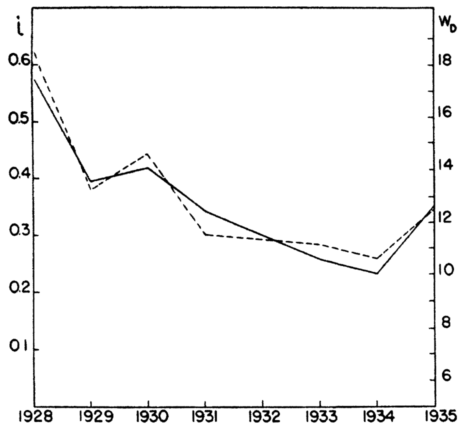

FIG. 17. Ratio i of H_e to H in solar prominences and magnetic disturbance W_D at Pavlovsk.[73]

The statistics of magnetic storms[75, 76] have shown that disturbed conditions tended to recur after intervals of about 27 days, and to explain this Maunder[74] suggested that the solar emission came from a restricted area of the solar surface in a narrow beam which rotated with the sun. The synodic rotation period of sunspots is 26.9, 27.3 and 28.3 days at heliographic latitudes 0°, 15°, and 30°, respectively; the periods of the reversing layer are about 2 percent less. Therefore, if the period of the beam were that of the solar surface the beam must come from a solar latitude of about 15°. On the assumption that the characteristics of a magnetic storm were somehow directly connected with the beam, the width of the beam is about 13° for a storm lasting one day and 1° for a 2 hour disturbance, provided the emission traveled approximately with the velocity of light; if the velocity were less the beam width is less. A sudden commencement called for a beam of exceedingly sharp forward edge. These requirements of the beam were stringent and evoked a shining remark from Chree,[77] "The theory advanced hardly touches the physical side of the problem, but it is clear and definite so far as it goes, and is obscured by no mystifications of language. It is certainly incomplete, and may be wholly erroneous, but it is exceedingly suggestive. . . ."

[75] Chree and Stagg, Phil. Trans. Roy. Soc. **A227**, 21 (1927); **212**, 75 (1912).
[76] Bartels, Terr. Mag. and Atmos. Elect. **37**, 1 (1932).
[77] Chree, Terr. Mag. and Atmos. Elect. **10**, 14 (1905).

The rotating narrow beam theory has met with difficulties in several directions. The surface rotation of the sun increases by about 6 percent from minimum to maximum of sunspots,[78] whereas no systematic lengthening or shortening of the 27 day recurrence interval could be detected for groups of years of many or few sunspots or for high or low sunspot latitudes.[75]

Terrestrial magnetic activity has long been recognized to undergo a seasonal variation according to a cosine law approximately with maxima at equinox and minima at solstice, the storm frequencies being about twice as great at the maximum as at the minimum.[79] The narrow beam theory sought to explain the seasonal changes by the tilt of the solar axis with respect to the earth, or in other words by the seasonal variation of the heliographic latitude of the earth.[80] Three separate rigorous tests of the idea by Bartels[76] all led to failure; his paper should be consulted for details. Maris[79] carried out detailed calculations based on the wide angle ultraviolet flare theory and was led to a reasonable explanation of the seasonal magnetic effect based on the tilt of the earth's axis with respect to the sun. At the same time he pointed out that probably all terrestrial magnetic disturbances are not of solar origin, but that some may arise from the passage of the earth through debris of planetary space.

Maris and Hulburt[81] looked to comets for evidence bearing on the width of the solar emission. They argued that if the same emission which caused the terrestrial magnetic storm and aurora fell on a comet it might cause changes in the comet. Many cases of apparently lawless and erratic behavior of comets, some famous in comet history, have been observed such as a sudden brightening, a splitting into pieces, emission of a new tail, etc. It was found that in general the comet events followed closely after a strong magnetic disturbance. A plot of the positions of the sun, earth, and comet at the epoque of 28 comet events showed that the

[78] Newall, Mon. Not. Roy. Ast. Soc. **82**, 101 (1921); Halm, Mon. Not. Roy. Ast. Soc. **82**, 479 (1922).
[79] Maris, Phys. Rev. **39**, 504 (1932).
[80] Cortie, Mon. Not. Roy. Ast. Soc. **73**, 52 (1913); Chapman, Mon. Not. Roy. Ast. Soc. **89**, 465 (1929).
[81] Maris and Hulburt, Phys. Rev. **33**, 1046 (1929).

average angle of the solar flare was about 90°. A direct association of comet events and terrestrial magnetic perturbation was incompatible with the narrow beam theory.

It is concluded that what evidence there is supports the wide angle flare hypothesis and is unfavorable to the narrow beam hypothesis. The conclusion leaves the recurrence tendency of magnetic storms without an explanation. This may be provided by the *ad hoc* assumption that the flare puffs out with the recurrence period due to solar pulsations of some sort.[81] Thus the two views are equally novel in that they call for unanticipated characteristics of the solar outburst, the rotating beam theory requiring a physically strange narrow beam and the wide puff theory a recurrent geyser action, which by coincidence has a period about the same as that of the rotation of the sun.

Conclusions

Theories of terrestrial magnetic variations and the aurorae are summarized and contrasted.

The theories of the solar diurnal variation of the terrestrial magnetic field are the dynamo, the diamagnetic and the drift current theories. The dynamo theory requires an upper atmospheric conductivity greater than that at first sight indicated by present inospheric measurements but not definitely denied by the measurements. In addition, it requires a general world-wide circulation of the high atmosphere, as would be produced by horizontal winds blowing away in all directions from regions under the sun. Such a wind system appears reasonable but is not known to exist. The diamagnetic theory likewise requires an ionization greater than that indicated but not definitely precluded by ionospheric observation. It is simpler than the dynamo theory in that it does not require world-wide winds. The drift current theory was shown to require a distribution of ionization over the earth in conflict with that calculated from the hypothesis that the cause of the ionization is the ultraviolet light of the sun. Ionosphere data are in accord with the hypothesis.

The only theory of the lunar variation of terrestrial magnetism is the dynamo theory which ascribed the lunar effect to assumed lunar tidal movements of the high atmosphere. Calculations based on sea-level barometric data yielded agreement with the general type of the observed lunar magnetic variation but of discordant phase. The discrepancy appears serious and has received no satisfactory explanation.

Theories of magnetic storms and aurorae are based on the assumptions that the effects are due to streams of charged particles or ultraviolet light from the sun. No complete charged particle theory adequate to account for the observed facts has been proposed. Various partial theories have been shown to be confronted with fundamental difficulties which they have never got around, in that the streams of particles cannot carry sufficient excess charge to yield the desired bending in the terrestrial magnetic field without at the same time giving rise to forces of electrostatic repulsion sufficient to disrupt the stream during its passage from the sun to the earth. The ultraviolet light theory appears adequate to explain the complicated storm variations of the terrestrial magnetic field, the auroral zone, the spreading of the aurora into low latitudes during strong storms, the delay of appearance of aurorae after incipience or the storm, and other observed facts. Recent ionosphere, radio and solar eruption data have proved that in certain cases the emanation from the sun which causes the terrestrial effects travels with the velocity of light within the error of observation, and thus offer general support of the ultraviolet light hypothesis.

The recurrence of magnetic disturbance at approximately twenty-seven day intervals may be explained by the hypothesis that the emanation causing the disturbance is emitted by the sun in narrow, less than 10°, beams which rotate with the sun or by wide angle, 90°, puffs with the recurrence period. The theories are equally novel in that they call for unanticipated characteristics of the solar emission, the beam theory requiring a physically strange narrow beam and the puff theory a recurrent geyser action or pulsation. Terrestrial and cometary evidence is adduced favorable to the puff theory and unfavorable to the beam theory.